

UNIVERSIDAD DE MÁLAGA
FACULTAD DE CIENCIAS DE LA EDUCACIÓN



TESIS DOCTORAL

**Evaluación de la competencia léxica bilingüe en
estudiantes del Máster Universitario en Profesorado**

Análisis de pruebas de disponibilidad léxica y de identificación de
tecnicismos en español, inglés y francés

Doctoranda: Inmaculada Clotilde Santos Díaz

Directores: Juan Andrés Villena Ponsoda y Antonio Manuel Ávila Muñoz

Programa Oficial de Postgrado en Políticas Educativas

Málaga, 2015



Publicaciones y
Divulgación Científica

AUTOR: Inmaculada Clotilde Santos Díaz

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está sujeta a una licencia Creative Commons:

Reconocimiento - No comercial - SinObraDerivada (cc-by-nc-nd):

[Http://creativecommons.org/licenses/by-nc-nd/3.0/es](http://creativecommons.org/licenses/by-nc-nd/3.0/es)

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer
obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de
Málaga (RIUMA): riuma.uma.es

A mis padres, que con su amor me dieron la vida
A mi marido, que con su amor me alegra la vida
A las personas a las que dedico los agradecimientos

Agradecimientos

El interés por la lingüística unida a mi vocación docente, me llevaron a delimitar el tema de esta tesis doctoral en el curso 2010/2011. Durante estos años, numerosas personas han contribuido de diversa forma en la realización de este trabajo. En primer lugar, cabe mencionar el asesoramiento de mis dos directores de tesis, Juan Andrés Villena Ponsoda y Antonio Ávila Muñoz. Ambos me han orientado con sabiduría y han desempeñado un papel relevante en el presente estudio. Por un lado, Juan me ha guiado por el apasionante y complejo mundo de la estadística donde he ido aprendiendo los conceptos y operaciones principales. Por otro lado, Antonio me abrió el camino en este campo de investigación hace casi 10 años cuando estudiaba la licenciatura en Traducción e Interpretación.

Al igual que la disponibilidad léxica, esta tesis ha viajado a los dos lados del Atlántico. En Holanda, me acogió el profesor Frans Hinkens durante una estancia en la que tuve la oportunidad de conocer el centro de investigación, Meertens Instituut. En Chile, la profesora Natalia Castillo Fadić me brindó su inestimable ayuda y cariño y me dio un gran impulso para acabar la tesis. Además, tuve la suerte de conocer a Alejandro Troncoso y a Roberto Ferreira y de poder realizarles consultas académicas.

Los datos del estudio han sido recabado gracias a la coordinación del profesorado del Máster Universitario en Profesorado y de la disponibilidad de su alumnado. Espero poder con los resultados de este estudio aplicar el principio de gratitud de Wolfram (1993) y devolver ese favor lingüístico.

Asimismo, quisiera expresar mi gratitud a todas las personas presentes en la bibliografía porque han aportado su granito de arena en la confección de esta tesis. En este sentido, agradezco el trabajo de los usuarios que cuelgan de forma anónima y desinteresada vídeos tutoriales y que tanto me han enseñado de estadística.

Por último, y no por ello menos importante, quisiera dar las gracias a mi familia (Santos Díaz y Conejo Sobrino), amigos y compañeros de trabajo que me han apoyado durante todo este tiempo y me han alentado en los momentos más difíciles. Mi mayor fortuna es poder contar con cada uno de ellos.

Gracias a todos por hacer posible que esta ilusión se haga realidad.

ÍNDICE

RÉSUMÉ	5
Chapitre 1. Introduction	7
Chapitre 2. Cadre théorique	9
Chapitre 3. Cadre méthodologique	11
Chapitre 4. Analyse générale	13
Chapitre 5. Analyse sociolinguistique	17
Chapitre 6. Proposition didactique	20
Chapitre 7. Conclusions	22
CAPÍTULO 1. INTRODUCCIÓN	27
CAPÍTULO 2. MARCO TEÓRICO	43
2.1. Introducción	47
2.2. Orígenes de la disponibilidad léxica	47
2.3. Principios epistemológicos y conceptuales de la disponibilidad léxica	57
2.4. Investigaciones relacionadas con el léxico disponible	76
2.4.1. En español como lengua materna	77
2.4.2. En inglés como lengua extranjera	97
2.5. Revisión del modelo matemático	101
2.5.1. Cálculo del índice de disponibilidad léxica	101
2.5.2. Índice de compatibilidad léxica y selección del léxico nuclear	111
2.5.3. Índice corregido de amplitud léxica	116
2.5.4. Índice corregido de descentralización léxica	118
2.5.5. Índice corregido de identificación de tecnicismos	119
2.6. Aplicaciones de la disponibilidad léxica	123
2.7. Variables que inciden en el léxico disponible	129
2.7.1. Sexo	132
2.7.2. Estancias en el extranjero	133
2.7.3. Conocimiento de idiomas	137
2.7.4. Corrección ortográfica	143
2.7.5. Clase social	145
2.7.6. Hábitos de lectura y exposición a los medios de comunicación	146
2.7.7. Contacto y uso de la lengua extranjera en diferentes contextos	147
2.8. Conclusiones parciales	150
CAPÍTULO 3. MARCO METODOLÓGICO	153
3.1. Introducción	157
3.2. Naturaleza y método de la investigación	158
3.3. Selección de la muestra	162
3.3.1. Tamaño de la muestra	163
3.3.2. Acceso a los informantes	166
3.3.3. Requisitos de acceso al Máster Universitario en Profesorado	168
3.3.4. Programa de estudios	170
3.4. Sistema de recogida de los datos	173
3.4.1. Encuesta sociológica	174
3.4.2. Prueba de disponibilidad	177
3.4.3. Prueba del léxico especializado	180
3.5. Edición y codificación de los datos	181
3.5.1. Encuesta sociológica	182
3.5.2. Prueba de disponibilidad	184

3.5.3. Prueba del léxico especializado.....	204
3.6. Conclusiones parciales.....	205
CAPÍTULO 4. ANÁLISIS GENERAL.....	209
4.1. Introducción.....	213
4.2. Análisis cuantitativo de palabras y vocablos.....	214
4.2.1. Palabras y vocablos en español.....	214
4.2.2. Palabras y vocablos en lengua extranjera.....	217
4.2.3. Palabras y vocablos en lengua inglesa.....	219
4.2.4. Palabras y vocablos en lengua francesa.....	221
4.3. Índice de densidad léxica e índice de cohesión.....	223
4.3.1. Índice de densidad léxica e índice de cohesión en español.....	224
4.3.2. Índice de densidad léxica e índice de cohesión en inglés.....	226
4.3.3. Índice de densidad léxica e índice de cohesión en francés.....	227
4.4. Promedio de palabras.....	227
4.4.1. Comparación intramuestral.....	228
4.4.2. Comparación intermuestral.....	233
4.5. Análisis de los vocablos con mayor índice de disponibilidad.....	237
4.5.1. Comparación intramuestral.....	238
4.5.2. Comparación intermuestral.....	242
4.6. La situación ortográfica.....	253
4.6.1. Resultados del estudio actual.....	253
4.6.2. Resultados comparativos intermuestrales.....	261
4.7. Incidencia de la variable 'sexo' en el léxico disponible.....	264
4.7.1. Centro 2: <i>La ropa</i>	269
4.7.2. Centro 3: <i>Comidas y bebidas</i>	270
4.7.3. Centro 4: <i>La escuela: muebles y materiales</i>	272
4.8. Conclusiones parciales.....	274
CAPÍTULO 5. ANÁLISIS SOCIOLINGÜÍSTICO.....	279
5.1. Introducción.....	283
5.2. Hipótesis.....	284
5.3. Análisis descriptivo univariante y presentación de las variables.....	287
5.3.1. Variables lingüísticas.....	292
5.3.2. Variables sociodemográficas y reticulares.....	295
5.3.3. Variables relacionadas con el dominio de una lengua.....	301
5.3.4. Resumen de las variables.....	309
5.4. Análisis bivariante.....	311
5.4.1. Relaciones entre variables lingüísticas.....	312
5.4.2. Efecto de las variables socioculturales.....	320
5.4.3. Efecto de las variables relacionadas con el dominio de una lengua.....	324
5.5. Análisis multivariante.....	346
5.5.1. Análisis de componentes principales (ACP). Reducción de dimensiones.....	347
5.5.2. Análisis de regresión lineal.....	356
5.5.3. Análisis causales.....	379
5.6. Conclusiones parciales.....	393
CAPÍTULO 6. SELECCIÓN DEL LÉXICO Y PROPUESTA METODOLÓGICA.....	397
6.1. Introducción.....	401
6.2. Organización de la memoria semántica.....	402

6.3. La teoría de los prototipos en los estudios de disponibilidad léxica.....	406
6.4. Procesos de reentrada del léxico.....	412
6.4.1. Modelos de reentrada según la ontogenia de los centros de interés	417
6.4.2. Modelo de reentrada de los centros de interés prototípicos	427
6.4.3. Modelo de reentrada de los centros de interés relacionales	440
6.4.4. Ejemplo de clasificación del centro de interés <i>Ordenadores e internet</i> según los procesos de reentrada.....	446
6.5. Análisis de las relaciones semánticas a través del DispoGrafo	452
6.5.1. Análisis de las relaciones semánticas según el centro de interés.....	453
6.5.2. Organización del léxico según el idioma	458
6.6. Hacia una propuesta didáctica	471
6.6.1. Selección léxica según diferentes criterios de corte	473
6.6.2. Los aspectos socioculturales en el léxico disponible	480
6.6.3. El léxico nuclear y el nivel según el Marco Común Europeo de Referencia para las Lenguas	483
6.6.4. Léxico nuclear en el Plan Curricular del Instituto Cervantes.....	494
6.6.5. Propuesta didáctica	498
6.7. Conclusiones parciales.....	501
CAPÍTULO 7. CONCLUSIONES.....	509
7.1. Conclusiones	513
7.1.1. Ámbito de la sociolingüística	513
7.1.2. Ámbito de la psicolingüística	523
7.1.3. Ámbito de la lingüística aplicada a la enseñanza de idiomas.....	529
7.1.4. Propuestas y líneas de investigación futuras	531
7.2. Conclusions.....	537
7.2.1. Cadre de la sociolinguistique.....	537
7.2.2. Cadre de la psycholinguistique	547
7.2.3. Domaine de la linguistique appliquée à l'enseignement des langues.....	552
7.2.4. Propositions et lignes de recherche futures	554
BIBLIOGRAFÍA	561
ANEXOS	613
Anexo I. Cuadro resumen de los objetivos, hipótesis y resultados de la tesis doctoral.....	615
Anexo II. Índice de gráficos, grafos y tablas	619
Anexo III. Documento de negociación	629
Anexo IV. Cuestionario sociológico y reticular	633
Anexo V. Pruebas de identificación de léxico especializado: español, inglés y francés.....	641
Anexo VI. Índice de siglas y abreviaturas	645
Anexo VII. Clasificación de errores ortográficos en español, inglés y francés	647
Anexo VIII. Selección de vocablos nucleares en español, inglés y francés.....	653
Anexo IX. Diccionarios de disponibilidad léxica en español, inglés y francés....	663

RÉSUMÉ

**Évaluation de la compétence lexicale bilingue des étudiants du
Master Universitaire en Professorat**

Analyse d'épreuves de disponibilité lexicale et du lexique spécialisé en
espagnol, anglais et français

Chapitre 1. Introduction

Ce chapitre se propose d'établir les fondements généraux de l'exposé et de l'étude traités plus en profondeur par la suite. Avant tout, il introduit et délimite le sujet en posant la question d'investigation. Ensuite, il présente les objectifs de la recherche à partir de différentes approches pour répondre à toutes les questions d'investigation. Ensuite, ce travail aspire à fournir une réponse à la question d'investigation par le biais d'hypothèses. Pour finir, il inclut la macrostructure du travail laquelle est divisée en sept chapitres.

Cette thèse de doctorat dirige son attention sur les facteurs qui conditionnent le vocabulaire virtuel en espagnol et en langue étrangère des étudiants de Master Universitaire en Professorat de l'Enseignement. Il se focalise sur le profil linguistique et sociologique des futurs professeurs et il a pour but d'analyser les types de relations entre les variables liées aux deux domaines. En outre, ce travail essaie d'évaluer le lexique disponible d'un des piliers basiques de l'implantation en Andalousie du *Plan de Fomento de Plurilingüismo*¹ : la formation initiale en langue espagnole et en langue étrangère du futur professorat. Étant donné les nombreux aspects que l'on peut étudier dans une recherche sur la formation linguistique des professeurs, cette étude sera limitée aux objectifs suivants:

1. Cadre de la sociolinguistique :

- Comparer le lexique disponible avec d'autres études pour détecter des ressemblances et des différences en termes de lieu, de temps, du niveau de formation des étudiants, de la situation de la langue (maternelle, étrangère, seconde langue).
- Analyser les variables qui conditionnent le lexique disponible en espagnol et dans la langue étrangère du futur professorat pour en extraire des données qui seraient d'intérêt dans l'enseignement des langues.

¹ Plan du Développement du Plurilinguisme adopté en 2005 selon l'Accord du 22 mars 2005 du Conseil de Gouvernement.

- Connaître le domaine du lexique spécialisé en espagnol et en langue étrangère selon les neuf centres d'intérêt du dictionnaire de disponibilité lexicale et chercher les facteurs qui ont une incidence plus ou moins importante.

2. Cadre de la psycholinguistique :

- Évaluer les connections de mots qui se produisent dans l'esprit du locuteur à travers le programme DispoGrafo (Echeverría, Vargas, Urzúa et Ferreira, 2008).
- Procéder à une analyse des méthodes de réentrée du lexique au noyau ou prototype de chaque centre d'intérêt par le biais de l'indice de compatibilité.
- Ébaucher un classement des centres d'intérêt selon différents paramètres.
- Confronter les analyses antérieures en langue espagnole à celles de la langue étrangère pour tirer des conclusions sur le lexique disponible au niveau quantitatif et qualitatif qui comprend l'organisation des mots.

3. Cadre de la linguistique appliquée à l'enseignement des langues :

- Élaborer un dictionnaire de disponibilité lexicale en langue espagnole et en langue étrangère (français et anglais) sur neuf centres d'intérêt de grande importance dans le domaine éducatif.
- Évaluer l'orthographe des mots et les erreurs les plus communes commises par les étudiants pour proposer des mesures correctives.
- Sélectionner les unités lexicales qui feront partie d'un vocabulaire de base à partir des listes du lexique disponible.
- Mener à bien une proposition didactique basée sur les résultats d'analyses réalisées, surtout sur celles liées à l'organisation des mots dans l'esprit.

Afin de répondre aux questions que pose cette étude, on propose des hypothèses comme élément de connexion entre la théorie et l'observation. De plus, les hypothèses ont guidé l'élaboration du questionnaire sociologique ainsi que l'épreuve de disponibilité lexicale et du lexique spécialisé. De cette façon, il est possible de corroborer ou réfuter les mêmes hypothèses à partir d'analyses qualitatives et quantitatives. Afin de suivre les objectifs, les hypothèses se divisent en trois blocs : cadre de la sociolinguistique, cadre de la psycholinguistique et cadre de la linguistique appliquée à l'enseignement des langues. Au fil des

chapitres, on expose les hypothèses et leurs résultats. Il est également possible de consulter les objectifs, hypothèses et conclusions dans l'Annexe I.

Chapitre 2. Cadre théorique

Ce chapitre analyse l'état de la question en allant du général au particulier. En premier lieu, on exposera la littérature rapportée à la disponibilité lexicale qui commence par un parcours des études pionnières. *Le Français Fondamental* est l'ouvrage de référence dans le domaine de la disponibilité lexicale. Après la chute de la IV République et la fin de l'Union Française, les objectifs de ce vocabulaire s'adaptent à la nouvelle situation : enseignement scolaire, diffusion de la langue dans les états qui viennent d'accéder à l'indépendance et où le français bénéficiait d'une situation privilégiée (ou était la langue officielle) et réponse aux besoins des touristes, étudiants et travailleurs qui se déplacent à l'étranger.

Ensuite, on parle de l'évolution de la méthodologie d'analyse et de la codification des données depuis l'ouvrage primitif, Gougenheim, Michéa, Rivenc et Sauvageot (1964) jusqu'à la consolidation des épreuves de listes ouvertes d'une durée de deux minutes établie par l'intérêt suscité par un projet commun entre l'Espagne et l'Amérique Hispanique dénommé *Proyecto Panhispánico de Disponibilidad Léxica*. Les chercheurs français (Gougenheim *et al.*, 1956, 1964) ont rassemblé des données de 904 scolaires de quatre départements : Dordogne, Marne, Eure et Vendée. Dans l'épreuve, chaque personne sondée devait écrire vingt unités lexicales liées à un centre d'intérêt donné jusqu'à un total de seize centres.

Cette méthodologie a été modifiée par Dimitrijević (1969) qui introduit la collecte des données par un système de listes ouvertes où les sujets fournissent des mots pour chaque centre dans une limite de temps précis. Les critères méthodologiques les plus diffusés dans ce domaine font leur apparition à la suite du Proyecto Panhispánico dans les années 90. La collecte des données se réalise conformément au système de listes ouvertes avec une durée de deux minutes pour chaque centre traditionnel présent dans l'ouvrage français. Le but du Panhispánico consiste à élaborer des dictionnaires de disponibilité lexicale dans les zones de langue

espagnole afin de réaliser des comparaisons dialectiques, sociolinguistiques, ethnolinguistiques, entre autres.

L'édition des données des épreuves de disponibilité lexicale est cruciale pour les analyses menées, surtout si l'intérêt réside dans la comparaison avec d'autres listes de mots. Samper Padilla (1998) établit la série de règles et de critères communs qui seront secondés par de nombreuses études, en particulier celles qui font partie du Panhispanico. Néanmoins, il y a des chercheurs qui plaident l'inclusion de tous les mots écrits par les locuteurs afin de connaître réellement le lexique disponible (Benítez Pérez, 1992b ; Mateo García, 1998 ; Galloso Camacho, 2002; Hernández Muñoz, 2004 ; Ávila Muñoz, 2006 ; Ávila Muñoz et Villena Ponsoda, 2010, 2011 ; etc.).

Dans ce sens, il est essentiel de connaître les recherches préalables sur le lexique disponible : plus particulièrement des recherches sur le lexique en espagnol comme langue maternelle et anglais et français comme langue étrangère. Dans le domaine hispanique, les enquêtes naissent du travail constant de López Morales qui publie en 1999 le lexique disponible de Porto Rico. Ensuite, les études commencent à apparaître en Amérique Hispanique comme en Espagne sur le lexique espagnol comme langue maternelle. Récemment, les études d'espagnol comme langue étrangère ou seconde langue, dans des pays hispanophones ou non, ont pris de l'importance : Carcedo González (2000c) en Finlande ; Moreno Fernández (2003) aux Etats-Unis ; González Fernández (2013) en Turquie, López González (2009) en Pologne, Bianca Sandu (2014) en Roumanie ; entre autres.

Le traitement mathématique des données fera l'objet de révisions où l'on exposera les principaux indices et les facteurs qui interviennent dans leur calcul. Il existe un consensus quasi généralisé pour calculer l'indice de disponibilité à travers de la formule de Strassburger et Chávez (1987) et l'indice de cohésion par le biais de l'équation proposée par Echeverría (1991). En outre, d'autres chercheurs, comme Ávila Muñoz et Villena Ponsoda (2010) complètent leurs analyses avec l'utilisation d'autres indices, tel que l'indice de compatibilité ou l'indice de décentralisation qui marquent le degré de ressemblance du lexique écrit par un informateur par rapport au reste des sujets. Cependant, García García (2014) réalise une révision mathématique

où il modifie l'indice de disponibilité, l'indice de compatibilité et l'indice de décentralisation. Il facilite aussi le calcul de ces indices à travers du programme SPSS. Avant, le calcul de l'indice de disponibilité et l'obtention des listes du lexique disponible se réalisait seulement avec un programme spécifique. Les applications les plus utilisées sont Lexidisp (Moreno Fernández, Moreno Fernández et García de las Heras, 1995), Dispolex (Bartol Hernández et Hernández Muñoz, 2004) et Dispogen (Echeverría Weasson, Urzúa Carmona, Figueroa, 2005).

Une fois abordés les aspects théoriques et méthodologiques, on a inclus les applications de la disponibilité dans plusieurs disciplines en mettant l'accent sur la linguistique appliquée à l'enseignement des langues. Le lexique qui en résulte fait l'objet de nombreuses analyses de différents domaines. Du point de vue de l'ethnolinguistique, la disponibilité permet de connaître la situation concrète d'une société à un moment précis ; d'après la sociolinguistique il est possible d'analyser l'incidence des facteurs sociaux dans le lexique à travers de variables sociales ; d'après la dialectologie on peut comparer les différentes syntopies d'une langue ; d'après la linguistique appliquée à l'enseignement des langues, il suppose être un outil très efficace pour la sélection du lexique ; et la psycholinguistique offre l'opportunité de connaître l'organisation des mots dans l'esprit.

Pour compléter ce chapitre, en suivent des recherches qui mettent l'accent sur les facteurs qui incident sur le lexique disponible en langue maternelle et en langue étrangère et qui seront analysés dans les chapitres suivants. Parmi ces variables, il faut souligner celles présentes dans la plupart des études, le sexe et la classe social ; d'autres rélevantes dans le lexique disponible en langue étrangère, la connaissance de langues et le niveau de compétence linguistique; variables peu étudiées, le nombre de livres et l'exposé aux moyens de communication et les fautes d'orthographe; et d'autres nouvelles dans ce domaine, les séjours à l'étranger.

Chapitre 3. Cadre méthodologique

Le but de ce chapitre est de présenter et de délimiter les fondements et méthodes suivis pour mener à bien cette étude. Une des premières décisions dans toutes les recherches est la délimitation de la population : universitaires de troisième

cycle qui étudient le Master en Professorat. De cette façon, il est possible de réaliser une évaluation lexicale des futurs professeurs. L'accès aux personnes sondées s'est réalisé par le biais d'un document de négociation qui a été préalablement distribué aux agents impliqués. On leur a communiqué les objectifs de la recherche et ses caractéristiques. En outre, on a remarqué la confidentialité des données et on a facilité les coordonnées de la candidate au doctorat aux professeurs et aux élèves dans le cas où ils sollicitent des références complémentaires.

L'échantillon de l'étude se replace dans son contexte puisqu'il est déterminant dans une recherche qui, entre autres, vise à servir d'analyse du lexique disponible du futur professorat. Pour ce faire, on montre le cadre normatif où se situe la formation initiale des professeurs et les conditions requises à l'accès au Master en Professorat. Comme couronnement, on spécifie le plan d'études que suivent les sondés qui font partie de l'univers d'étude à l'Université de Malaga.

Le travail des enseignants dans l'implantation du plan du plurilinguisme est crucial ainsi que leur formation initiale². Le professorat de l'école primaire doit avoir obtenu le diplôme professionnel de professeur des écoles (DPPE). Cependant, les professeurs d'enseignement secondaire et de formation professionnelle doivent certifier la réalisation d'une maîtrise et le Master Universitaire en Professorat de l'Enseignement Secondaire Obligatoire, Baccalauréat, Formation Professionnelle et Enseignement des Langues ³. Ce master substitue le Certificat d'Aptitude Pédagogique en vigueur jusqu'en 2008/2009 et où il n'était pas nécessaire l'accréditation d'un niveau B1 en langue étrangère.

Ensuite, on détaille l'élaboration des documents pour obtenir l'information. Les données extralinguistiques se sont recueillies à travers d'un questionnaire avec des questions sociologiques, réticulaires et liées aux procédés d'acquisition et apprentissage de la langue étrangère du sujet. L'information linguistique provient des épreuves de disponibilité lexicale et d'un test d'identification de mots techniques. Les centres d'intérêt en espagnol et en langue étrangère (anglais ou français) sont neuf :

² Il faut signaler que dans les années 2013/2014 et 2014/2015 les bourses d'accès au travail d'enseignants bilingues se sont ouvertes au personnel intérimaire puisqu'il n'y avait pas assez de professeurs avec l'accréditation d'un niveau B2 en anglais, français et allemand.

³ Traduction de la dénomination donnée au master à l'Université de Malaga.

huit correspondent aux centres traditionnels⁴ présents dans l'étude de Gougenheim *et al.* (1964) et un centre de nouvelle incorporation, *Ordinateurs et internet*.

Une fois les données compilées, on a réalisé un plan de codification pour le questionnaire sociologique où on attribue une valeur numérique à chaque réponse qui s'incorpore à une base de données. Pour l'édition du lexique disponible, on a suivi les recommandations de Samper Padilla (1998) soutenues par la plupart des travaux héritiers du Panhispanico. En plus, on a introduit des critères propres, tels que l'incorporation aux listes des nombres propres et les associations individuelles qui permettent de réaliser des analyses sur la relation entre les mots plus fidèles. Les particularités de l'édition se sont présentées en détail selon la langue et le centre d'intérêt en question pour justifier les critères établis et orienter les futurs chercheurs. Quant à l'épreuve d'identification du lexique spécialisé, les réponses se sont classifiées en correctes, incorrectes et non répondues⁵.

En synthèse, cette étude se sert de la méthodologie de base des recherches de disponibilité qui tournent autour du Panhispanico (épreuve de deux minutes sur un centre d'intérêt, système de listes ouvertes, etc.) mais elle incorpore aussi des progrès méthodologiques, tel que l'épreuve d'identification du lexique spécialisé créée par Ávila Muñoz y Villena Ponsoda (2010) ou l'utilisation du programme DispoGrafo (Echeverría, Urzúa et Ferreira, 2008). En plus, on incorpore un échantillon original sur lequel on évalue le lexique en langue maternelle et en langue étrangère. Comme signale Moreno Fernández (1990: 38)⁶, la seule façon de faire des progrès est à travers l'union de la tradition et originalité.

Chapitre 4. Analyse générale

Les analyses suivantes montrent les données générales liées à l'épreuve de disponibilité lexicale : nombre total de mots par épreuve et par centre d'intérêt,

⁴ 1) *Les parties du corps humain* ; 2) *Les vêtements* ; 3) *La nourriture et les boissons* ; 4) *L'école (meubles et matériel)* ; 5) *La ville* ; 6) *Moyens de transport* ; 7) *Jeux et loisirs* ; 8) *Professions et métiers*.

⁵ Le procédé mathématique de ce test ainsi que l'épreuve du lexique disponible trouvent leurs fondements dans la section 2.5 et se mettent en pratique sous la forme d'analyses dans les chapitres 4, 5, et 6.

⁶ «Sólo así la comparación es posible, sólo así los pasos de una persona pueden ser de plena utilidad para otra, porque los avances se consiguen a mitad de camino entre la tradición y la originalidad» (Moreno Fernández, 1990 : 38).

nombre de vocables, indice de cohésion et densité lexicale. D'une part, on analyse le lexique en langue espagnole et en langue étrangère des futurs professeurs et, d'autre part, on compare le lexique en langue espagnole de l'étude actuelle avec de recherches préalables.

D'abord, on présente l'analyse du nombre de mots et vocables qui se complète à travers la comparaison de la moyenne des unités lexicales de cette étude par rapport à des recherches antérieures. La langue maternelle des sondés résulte être la plus productive quant au nombre de mots et vocables. En espagnol, les sujets ont écrit une moyenne de 191,91 mots, en anglais 131,79 et en français 147,71. En ce qui concerne les vocables, en espagnol on registre une moyenne de 20,93, en anglais de 13,77 et en français de 5,82. Ces données si différentes s'expliquent parce que le nombre de vocables est influencé par le nombre d'informateurs : 171 en espagnol, 150 en anglais y 21 en français.

La comparaison des résultats avec d'autres études réalisées à Malaga, particulièrement sur un échantillon d'étudiants de baccalauréat (Ávila Muñoz, 2006) et de population adulte (Ávila Muñoz et Villena Ponsoda, 2010), montre la plus grande productivité des universitaires de troisième cycle. La moyenne des mots par centre d'intérêt calculée sur les huit centres en commun, est de 19,69 chez les étudiants de baccalauréat, 20,50 de la population adulte stratifiée et 21,59 des étudiants universitaires de troisième cycle. *A priori*, le niveau de formation paraît avoir une certaine influence sur le lexique disponible. Dans les études de disponibilité lexicale en langue étrangère, on corrobore aussi l'hypothèse que la capacité lexicale en langue maternelle est supérieure à la langue étrangère.

En second lieu, on présente les résultats de l'indice de densité lexicale et l'indice de cohésion. Dans ce cas, on ne réalise pas de comparaison avec d'autres études parce qu'il s'agit de deux indices directement liés au nombre de sujets. On constate l'existence d'une tendance commune dans les trois langues de façon qu'il y ait des centres dont les rangs se répètent selon leur indice de cohésion et densité lexicale. Le centre le plus adhérent est le premier, *Parties du corps humain*, dans les trois langues tandis que les plus hétérogènes sont *Jeux et loisirs* et *Ordinateurs et internet*. Si l'on compare le classement d'autres études, comme celle d'Ávila Muñoz et Villena

Ponsoda (2010), on observe que le centre le plus homogène est le premier. C'est pourquoi dans le chapitre 6, on présente en profondeur la structure interne des centres d'intérêt se basant sur le fait que le type de centre conditionne les réponses des sujets.

En troisième lieu, on centre l'attention sur les vocables les plus disponibles afin d'analyser la différence entre différents échantillons et entre le lexique apporté en langue maternelle et en langue étrangère des informateurs. On a constaté l'hypothèse qui présuppose une majeure similitude dans les études d'une langue étrangère par rapport aux études en langue maternelle. D'une part, l'usage et le contexte d'apprentissage influencent directement sur le lexique disponible. De cette façon, parmi les vocables les plus disponibles en langue maternelle on trouve quatre liés aux sous-vêtements tandis qu'en langue étrangère, ceux-ci ont été écrits seulement une ou deux fois. Ce type de vêtements est utilisé tous les jours mais n'apparaît pas si fréquemment dans les manuels d'anglais, français ou d'espagnol comme langue étrangère. D'autre part, l'espace temporel de récolte des données montre la diachronie de la langue et conditionne le vocabulaire plus disponible. Pour cette raison, dans l'étude pionnière de Gougenheim *et al.* (1964) apparaissent des vocables tels que *combinaison* qui ne seraient pas disponibles de nos jours parce que l'on a changé notre façon de s'habiller.

Dans la section suivante, on présente les fautes d'orthographe des enquêtes⁷ en espagnol, en anglais et en français. D'abord, on compare les résultats des trois langues de cette étude et ensuite on réalise la comparaison avec des recherches préalables (García Paredes, 1999 ; Ávila Muñoz, 2007 ; Saura Rami, 2008 ; Blanco Domínguez, 2011 ; García Casero, 2013).

En espagnol, dans chaque groupe de 100 mots écrits par sujet, 1,32 contient une erreur tandis qu'en langue étrangère 4,90. Le 66,16 % des fautes d'orthographe sont dues à une inadéquate accentuation et le 3,95 % à la confusion des lettres *b* et *v*. Quant à la distribution des vocables avec des erreurs au respect du total des mots écrits, sauf exceptions, ils se correspondent aux centres avec un indice d'erreurs d'orthographe élevé, tel que le centre sept et neuf pour se trouver avec un nombre considérable de mots familiers et d'origine étrangère écrits de façon incorrecte.

⁷ Pour accéder aux fautes d'orthographe, voir l'Annexe VI.

Les universitaires de troisième cycle obtiennent des résultats considérablement meilleurs que les étudiants universitaires malaguènes, les élèves de lycée aragonais, galiciens et madrilènes. Parmi les étudiants de baccalauréat, les galiciens occupent le premier poste en correction orthographique (4,59 % des fautes), suivis des aragonais (7,72 %) et des madrilènes (11,34 %). En langue étrangère, les sujets de cette étude ont commis un 4,9 % de fautes et se situent seulement après les galiciens. Pour cette raison, on peut conclure qu'il a un certain progrès dans la correction orthographique des étudiants acquise tout au long des études universitaires.

Du point de vue de la sociolinguistique, on étudie l'effet du sociolecte féminin et masculin dans le nombre de mots et de vocables apportés par les sondés. Pour ce faire, on a réalisé une analyse de corrélation bivariée et on a mis le point sur les centres qui présentent une différence statistiquement significative. Le second centre, *Les vêtements*, présente une corrélation significative bilatérale à niveau 0,01 et le troisième et quatrième centre, *La nourriture et les boissons* et *L'école : meubles et matériel scolaire* à niveau 0,05.

De façon plus détaillée, on a comptabilisé et commenté le lexique en commun présent dans les listes qui contiennent les 20 vocables plus disponibles des femmes et des hommes pour analyser d'un point de vue qualitatif les différences. Dans ce sens, on a distingué les vocables apportés de manière exclusive par les hommes et les femmes et on a constaté qu'il existent certaines différences. Par exemple, dans le centre dénommé *La nourriture et les boissons* les hommes ont l'habitude d'écrire plus de noms liés aux boissons alcooliques tandis que les femmes ont mentionné plus de vocables en relation aux desserts et aliments sucrés.

En ce qui concerne le nombre de vocables par sujet et centre d'intérêt, la moyenne par centre d'intérêt des femmes est supérieure dans tous les centres à l'exception du cinquième et neuvième, *La ville* et *Ordinateurs et Internet* en langue espagnole et le sixième et neuvième en langue étrangère *Moyens de transport* et *Ordinateurs et Internet*. Cette analyse corrobore l'hypothèse de Bellón Fernández (2011 : 119) qui affirme que le centre d'intérêt marque la différence entre un sexe et l'autre puisqu'il y a certains centres où le sexe féminin est plus enclin à écrire plus de mots et vice versa.

Les conclusions de cette section coïncident avec les études préalables où les différences entre le sociolecte féminin et masculin sont très faibles (Gougenheim *et al.* 1956 ; Dimitrijević, 1969 ; Echeverría, Herrera, Moreno et Pradenas, 1987; Alba, 1995a ; Mateo, 1998 ; Valencia et Echeverría, 1991 ; Carcedo, 2001). Cependant, de même que dans l'étude de Valencia et Echeverría (1999 : 257), on constate que dans la plupart des centres d'intérêt, le sexe féminin est plus productif.

Chapitre 5. Analyse sociolinguistique

Tout au long de ce chapitre on essaiera de connaître quels sont les facteurs qui ont une incidence sur le lexique disponible et sur la capacité d'identification des mots techniques en espagnol et en langue étrangère. Le premier pas pour réaliser une estimation optimale des résultats passe par une étude de ses variables de manière individuelle. À travers de l'analyse univariée, on prétend savoir la distribution et dispersion de chaque variable isolée. Cette phase est indispensable comme approximation aux analyses bivariées et multivariées présentées ultérieurement. Chacune de ces variables est une variable échelle ou univariée et l'ensemble des variables font partie d'une variable vectorielle ou multivariée.

Les variables les plus significatives de l'étude sont divisées en trois types : variables sociodémographiques et réticulaires, variables liées à l'apprentissage d'une langue étrangère et variables linguistiques. Les deux premières font référence aux variables indépendantes qui aident à expliquer le comportement des troisièmes, c'est-à-dire, des variables dépendantes. Il faut souligner que la base des données primitive est composée de plus de 500 variables comme résultat de la codification du questionnaire sociologique et des épreuves linguistiques mais on a choisi seulement les variables les plus significatives.

Parmi les variables sociodémographiques et réticulaires, on a inclus l'âge de la personne sondée, les études et le travail des parents et les revenus familiaux. Quant aux variables liées à l'apprentissage d'une langue étrangère, on expose les données exploratoires plus remarquables : spécialité du master, utilisation de la langue étrangère dans son environnement, exposition aux moyens de communication et la motivation pour donner des cours en langue étrangère. Par rapport aux variables

linguistiques, on présente celles qui seront utilisées comme variables dépendantes dans les analyses bivariées et multivariées : l'indice corrigé de décentralisation lexicale (ICoDD)⁸ et l'indice corrigé d'identification de mots techniques ICIMT)⁹ en espagnol et en langue étrangère.

L'indice corrigé de décentralisation lexicale mesure la capacité lexicale des sondés à partir des résultats de l'épreuve de disponibilité lexicale. Pour son calcul, on a considéré les mots moins compatibles ou mentionnés seulement par une minorité de sujets et la relation avec le centre d'intérêt. De cette façon, on obtient un indice qui fait référence au nombre de mots par locuteur et aussi au degré de décentralisation par rapport au centre d'intérêt.

En langue étrangère les facteurs qui gardent une plus grande relation avec l'ICoDD sont liés en grande partie à l'apprentissage d'une langue étrangère et au propre ICoDD en langue espagnole. Dans le modèle de régression linéale et le modèle causale on a remarqué la motivation pour donner des cours en enseignement bilingue, l'usage de la langue étrangère dans son contexte naturel et la spécialité du master qui l'étudient les sondés. Ils existent d'autres variables, telles que la connaissance d'autres langues étrangères, qui présentent une relation significative avec l'ICoDD mais qui sont redondantes dans les modèles multivariés parce qu'elles sont liées surtout à la spécialité du master.

En espagnol les analyses bivariées ainsi que les multivariées montrent une relation moins significative entre les variables linguistiques et extralinguistiques avec l'ICoDD. En outre, l'analyse de régression linéale et l'analyse causale expliquent avec un degré d'adéquation plus petit qu'en langue étrangère le comportement de la capacité lexicale en espagnol. La principale cause de cet effet est due au rôle principal

⁸ Cette variable permet de connaître le degré de spécificité des mots apportés par un informateur par rapport aux unités lexicales de l'ensemble de l'échantillon. L'indice de décentralisation comptabilise tous les mots avec la valeur 1 sans tenir compte de la pertinence plus ou moins directe du vocable avec le centre d'intérêt. Cependant, Ávila Muñoz et Villena Ponsoda (2010 : 185) proposent de corriger cet indice et appliquent des valeurs différentes selon le type de relation des vocables par rapport au centre d'intérêt ou noyau prototypique : mots nucléaires (N1 = 1), associations collectives (N5 = 0,5), dérivations (N25 = 0,25) et associations individuelles (N0 = 0). La somme des mots préalablement pondérée selon ce classement donne lieu à l'indice corrigé de décentralisation lexicale.

⁹ Dans l'étude d'Ávila Muñoz et Villena Ponsoda (2010 : 217) les résultats de l'épreuve de lexique spécialisé se sont montrés à travers le Pident, défini comme la capacité individuelle d'identification de pièces lexicales spécialisées. Cet indice s'obtient de la somme de toutes les réponses correctes des locuteurs sans tenir compte le hasard dans ce test. Pour corriger l'indice d'identification de mots techniques, on a soustrait les réponses incorrectes avec une valeur statistiquement pondérée.

que joue le degré de formation des locuteurs dans leurs compétences lexicales. Étant donné que tous les étudiants sont universitaires de troisième cycle, il existe une marge plus petite d'incidence dans la capacité lexicale qui vient marquée par d'autres variables linguistiques, l'ICoDD en langue étrangère et l'ICIMT en espagnol, ou directement liés à celles-ci, la spécialité du master et les habitudes de lecture.

En ce qui concerne l'indice corrigé d'identification de mots techniques (ICIMT) mesure la capacité de l'informateur de reconnaître des mots techniques et en relation avec le centre d'intérêt donné. Il s'obtient à partir d'une formule mathématique qui tient en compte des réponses correctes et incorrectes qui en résultent de l'épreuve du lexique spécialisé. Cette épreuve est centrée sur le vocabulaire passif du locuteur puisqu'elle vérifie si le sujet est capable de reconnaître certaines unités lexicales mais cela n'implique pas nécessairement que le sujet puisse utiliser le mot en contexte et qui fasse partie du vocabulaire actif.

Les analyses de régression linéale multiples et causales qui mesurent les facteurs qui influencent la capacité d'identification de mots spécialisés, présentent un modèle d'adéquation assez faible en langue étrangère et, surtout, en espagnol. Cependant, il est intéressant de souligner les variables avec une plus grande influence sur l'ICIMT en langue maternelle et étrangère pour les prendre en compte dans le domaine de la linguistique appliquée à l'enseignement des langues.

En langue étrangère les variables avec une incidence positive sur l'ICIMT sont l'usage de la langue dans de différents contextes, l'ICIMT en espagnol, l'élection du français dans l'épreuve, l'ICoDD en langue étrangère et la classe sociale. De même que dans les modèles multivariés de la capacité lexicale, les variables telles que la connaissance d'autres langues gardent une relation significative bilatérale avec l'ICIMT sont présents dans des modèles multivariés de façon indirecte, par exemple à travers l'utilisation de la langue et la spécialité du master.

En langue espagnole il se produit une relation d'influence réciproque avec l'ICIMT en langue étrangère. Le reste de variables remarquables sont les habitudes de lecture, l'exposé aux programmes culturels, la spécialité du master et la classe sociale. C'est pourquoi une des conclusions possibles du point de vue pragmatique et

didactique est que le nombre de livres qu'un étudiant lit par an a une influence sur le vocabulaire actif et passif.

Chapitre 6. Proposition didactique

Cette proposition vise à fournir aux professionnels du domaine de l'enseignement de l'espagnol comme langue maternelle ou langue étrangère des instruments pour sélectionner le lexique selon le niveau de leurs élèves. D'abord, il définit les fondements théoriques des principaux modèles d'organisations de la mémoire sémantique : théorie des traits, théorie des réseaux sémantiques, théorie des prototypes et postulés de signifié. Ensuite, deux modèles se développent selon Hernández Muñoz (2006: 67), ils sont plus attachés à la disponibilité lexicale. De cette façon, il est possible de montrer les relations entre des mots de deux manières différentes mais complémentaires.

D'un côté, on présente le lexique prototypique ou nucléaire généré à partir de l'application de la théorie des ensembles diffus aux résultats de l'indice de compatibilité lexicale. On montre le vocabulaire des listes en espagnol, anglais et français d'un centre d'intérêt pour comparer le degré de coïncidence entre le lexique de base¹⁰. Une fois défini le lexique de base, on évalue les procédés de réentrée au noyau à partir du modèle d'Ávila Muñoz et Sánchez Sáez (2010, 2011 et 2014). Il s'agit de mesurer la compatibilité de chaque vocable et de la présenter graphiquement selon l'ordre d'apparition de ceux-ci.

Dans le but de montrer que la différente ontogénie des centres d'intérêt influence sur l'organisation des mots, on a sélectionné un centre prototypique, *Parties du corps humain*, un centre relationnel, *L'école : meubles et matériel*, et le centre qui n'avait pas été classifié avant dans d'autres études, *Ordinateurs et internet*. Il faut préciser qu'après avoir analysé les taxonomies proposées par Hernández Muñoz (2006), Tomé Cornejo (2011) et Sánchez-Saus Laserna (2011) basées sur les aspects théoriques des centres d'intérêt, on a proposé un classement fruit des analyses des résultats.

¹⁰ Le reste des listes avec le lexique compatible de tous les centres d'intérêt dans les trois langues se trouvent dans l'Annexe VIII.

Le classement ordonne sous forme d'échelle les centres d'intérêt selon leur indice structurel. Cet indice se calcule de la somme de l'indice de cohésion, l'indice de centralité et l'indice de connectivité. L'objectif n'a pas été d'établir une dichotomie entre centres prototypiques et relationnels mais plutôt positionner les centres sur une ligne continue qui fournisse l'information qui provient de la pratique, c'est-à-dire, des résultats mêmes de l'épreuve de disponibilité lexicale.

D'un autre côté, on réalise une analyse des relations avec le programme DispoGrafo. L'objectif est de montrer de manière visuelle les liens entre les différents nœuds ou vocables et le poids ou la force de ces relations entre les trois langues d'étude. Les résultats permettront de connaître les relations de chaque unité lexicale et son degré de connectivité afin de rapprocher les propositions didactiques au réseau sémantique des locuteurs.

À l'instar de la théorie des prototypes, les relations entre les mots sont subordonnées au type de centre d'intérêt. Les centres prototypiques présentes des sous-champs lexicaux mieux définis et avec une plus grande importance. Néanmoins, les centres relationnels comptent sur un nombre plus grand de combinaisons entre des mots, ce qui entraîne une dispersion des relations.

Sur la même lignée que Ferreira (2006), les relations entre les mots des centres en langue étrangère ne sont pas si claires. Par contre, les sondés qui possèdent une compétence linguistique élevée en langue étrangère ont l'habitude de reproduire la séquence lexicale apportée en espagnol. Dans certains cas, la taille, c'est-à-dire, la sélection des relations selon leur poids, a été très faible puisque dans le cas contraire toutes les relations disparaissent par manque de fréquence donnée.

Pour finir, il faut tenir compte de ces résultats pour réfléchir à l'importance d'un programme qui permette de connaître quel est le vocabulaire le plus adéquat pour les étudiants selon leur compétence linguistique. Il s'agit d'abord de connaître la similitude des centres d'intérêt par rapport aux notions spécifiques du Plan Curriculaire de l'Instituto Cervantes pour évaluer la possible modification des centres traditionnels. Ensuite, on propose la création d'un corpus ou une base de données qui inclut les trois grands blocs : le lexique nucléaire, les relations entre les mots et la

correspondance de ce dernier avec le Cadre européen commun de référence pour les langues.

Chapitre 7. Conclusions

Les conclusions maintiennent une étroite relation avec les objectifs et hypothèses de l'étude. Pour cette raison, elles sont divisées selon les trois cadres qui recouvrent cette recherche : sociolinguistique, psycholinguistique et linguistique appliquée à l'enseignement des langues. Il faut souligner que chaque chapitre possède une section dénommée conclusions partielles où on introduit un résumé des résultats principaux.

a) Cadre de la sociolinguistique

- La moyenne des mots par informateurs des étudiants de baccalauréat de Málaga de l'étude d'Ávila Muñoz (2006) est de 19,69 tandis que les universitaires de troisième cycle ont apporté une moyenne de 21,59 pour les huit centres d'intérêt communs.
- En espagnol les étudiants ont écrit une moyenne de 191,91 mots, en anglais 131,79 et en français 147,71. Les comparaisons avec d'autres études (Carcedo González, 2000c ; Ferreira, 2006 ; entre autres) montrent qu'en langue maternelle, la moyenne de mots est toujours supérieure à celle de la langue étrangère.
- Dans la capacité lexicale en langue étrangère les variables avec une plus grande incidence sont : la spécialité du master, la capacité lexicale en espagnol, la motivation du futur professorat et l'utilisation de la langue étrangère dans des différents contextes.
- Dans la capacité lexicale en langue espagnole les variables les plus influentes sont : le niveau d'instruction, la capacité lexicale en langue étrangère et la spécialité du master¹¹ et la capacité d'identification des mots techniques.

¹¹ L'influence des spécialistes varie selon la langue étrangère: Mathématiques est la seule spécialité qui a une incidence négative. Langue Étrangère en Anglais et Français, Langue Castellane et Littérature, Latin et Grec et Musique ont une incidence positive et le groupe formé par les spécialités de Géographie et Histoire, Dessin,

- En langue étrangère, les variables avec une incidence positive sur l'indice corrigé d'identification de mots techniques (ICIMT) sont l'utilisation de la langue dans de différents contextes, l'ICIMT en espagnol, l'élection du français dans l'épreuve, l'ICoDD en langue étrangère et la classe sociale. La connaissance d'autres langues garde une relation significative bilatérale avec l'ICIMT qui est présent dans les modèles multivariés de manière indirecte, par exemple à travers l'utilisation de la langue et la spécialité du master.
- En langue espagnole se produit la relation d'influence réciproque avec l'ICIMT en langue étrangère. Les autres variables remarquables sont les habitudes de lecture, l'exposé aux programmes culturels, la spécialité du master et la classe sociale. C'est pourquoi une des conclusions que l'on peut extraire d'un point de vue pragmatique et didactique est que le nombre de livres qu'un étudiant lit par an influence son vocabulaire actif et passif.
- Les analyses statistiques montrent une relation significative à un niveau inférieur de 0,01 entre l'ICoDD et l'ICIMT. En outre, l'influence d'une variable sur l'autre est manifeste dans les analyses multivariées.

b) Cadre de la psycholinguistique

- Après avoir réalisé les analyses des procédés d'entrée au noyau et de la prototypicité des mots, on propose une taxonomie des centres d'intérêt. Chaque centre se situe sur un continuum qui part des centres prototypiques jusqu'aux centres relationnels. Pour les classer, on tient compte de la cohésion du centre, de la prototypicité du noyau et des relations entre les mots.
- Les modèles de réentrée au noyau varient en fonction de la capacité lexicale de l'informateur et du type de centre d'intérêt. Dans les centres considérés plus prototypiques on remarque trois types de procédés de réentrée : léger, modéré

et prononcé qui correspondent respectivement aux sujets ayant une faible, moyenne et haute capacité lexicale.

- Par rapport aux centres relationnels, on ne localise que deux modèles de réentrée : le modèle de réentrée modéré et le modèle prononcé avec une déviation linéale. Les locuteurs qui ont apporté peu de mots ne produisent pas de modèles de réentrée légers puisque le noyau n'est pas si défini et cela provoque que tous les mots ne présentent pas de niveaux si hauts de compatibilité.
- Les graphes montrent l'organisation des mots dans les sous-champs lexicaux, l'information qui se complète avec les données sur les relations entre les unités lexicales. Effectivement, les centres homogènes dessinent des réseaux bien plus définis et des relations plus fortes entre les mots.

c) Cadre de la linguistique appliquée à l'enseignement des langues :

- Les universitaires de troisième cycle commettent une moyenne de 0,28 fautes d'orthographe par sujet et centre d'intérêt et 0,73 en langue étrangère. Tandis que les élèves de baccalauréat madrilènes (Paredes García, 1999) montrent une moyenne de 2,26 erreurs et les aragonais 1,83 (Saura Rami, 2008). Le 66,16 % des fautes d'orthographe correspondent à une inadéquate accentuation des mots.
- À travers la théorie des ensembles diffuses appliquée par Ávila Muñoz et Sánchez Sáez (2010, 2011), il est possible de sélectionner le lexique de base. Ce vocabulaire correspond au niveau d'utilisateur élémentaire établi par le Cadre européen commun de référence pour les langues (CECRL). On a constaté le degré de similitude entre le lexique nucléaire en espagnol, anglais et français qui ratifie que les mots se situent autour d'un noyau ou prototype. C'est pourquoi cette méthode devient un outil clé pour rassembler des critères dans les études de disponibilité lexicale dans la sélection du vocabulaire de base qui correspond aux niveaux A1 y A2 selon le CECRL.

Pour finir, on ajoute des lignes d'investigation futures où on détaille la création d'une base de données ou corpus qui inclut les résultats de cette étude. Les différents fondements de cette proposition sont prouvés tout au long des sept

chapitres de façon à dévoiler de nouvelles voies de découverte. En outre, les résultats ouvrent la porte à des applications pratiques dans l'organisation des plans d'études, de systèmes de bourses à l'étranger, de l'encouragement des habitudes de lecture, entre autres. Dans ce sens, les comparaisons avec d'autres études de la même nature au niveau local ou international permettront de relever d'autres variables et de corroborer celles de cette recherche.

CAPÍTULO 1. INTRODUCCIÓN

A modo de prolegómenos, este capítulo pretende establecer los fundamentos generales que llevaron al planteamiento y realización del estudio actual y que serán tratados con mayor profundidad en los siguientes capítulos. En primer lugar, se introduce y limita el tema planteando el problema de investigación¹². A continuación, se presentan los objetivos que persigue la investigación desde enfoques diversos para conocer la respuesta a los interrogantes del estudio. Posteriormente, se intenta dar una respuesta tentativa al problema de investigación mediante las hipótesis. Por último, se incluye la macroestructura de la tesis que se divide en siete capítulos.

Los cambios en las estructuras sociales, en la producción y distribución de conocimiento científico, en los procesos económicos y cambios a nivel tecnológico y de desplazamientos humanos en busca de una mayor calidad de vida y un mayor bienestar en los últimos tiempos, están generando una sociedad más multicultural. La necesidad de comunicación en una lengua extranjera se ha convertido en un requisito *sine qua non* en el ámbito profesional, cultural, de ocio y de intercambio de información (Commission Européenne, 1995¹³; Guillén Díaz, 2000; Consejo de Europa, 2001; Consejería de Educación de la Junta de Andalucía, 2005). En la perspectiva de una enseñanza globalizadora, las lenguas extranjeras pueden contribuir de manera muy activa a desarrollar las capacidades cognitivas, afectivas, psicomotrices, de creatividad, de relación personal y de inserción social de los jóvenes. La lengua viva ofrece, en relación con otras asignaturas escolares, la ventaja de poder, sin perder su especificidad, actuar al mismo tiempo (por mimetismo, simpatía o empatía) sobre otros saberes y otros dominios de formación.

El Consejo de Europa insiste en la necesidad de desarrollar competencias que promuevan la comunicación de los individuos con sus vecinos europeos. Para hacer efectivo tal cometido, el Consejo publica en el año 2001 el Marco Común Europeo

¹² «Definir el problema de la investigación equivale a seleccionar una dirección concreta (...) o seleccionar un evento, una situación, un hecho, un comportamiento y delimitar el tiempo, el espacio, las personas, el contexto en donde uno se decide a investigar» (Ruiz Olabuénaga 2003: 54).

¹³ *Vid.* versión electrónica en el siguiente enlace: http://europa.eu/documents/comm/white_papers/pdf/com95_590_fr.pdf [Fecha de última consulta el 8 de febrero de 2015]. La página web oficial http://ec.europa.eu/white-papers/index_es.htm recoge los Libros Blancos publicados desde el año 1993. En esta ocasión el Libro blanco está disponible en inglés, francés y alemán, de ahí que se haya citado en francés. [Fecha de última consulta el 8 de mayo de 2015].

de Referencia para las Lenguas (MCERL). En general, las leyes educativas de los diferentes países europeos establecen, de acuerdo con las exigencias del Consejo de Europa, la enseñanza de dos lenguas extranjeras habladas en Europa durante la etapa de educación secundaria. Cabe destacar, que en la enseñanza primaria e incluso ya en la enseñanza infantil, se ha incluido la impartición de una lengua extranjera, preferentemente la lengua inglesa (excepto en los países anglófonos) considerada primera lengua extranjera mientras que la segunda lengua extranjera (que se empieza a enseñar *a posteriori*) varía según los acuerdos entre países, los intereses económicos y las necesidades sociales. Así, el papel de la lengua se ve reforzado a medida que aumenta el interés o necesidad de los individuos para comunicarse en inglés.

Con el objetivo de adaptar la sociedad andaluza a los nuevos retos que plantea el fenómeno de la globalización, el Gobierno Andaluz publicó el documento titulado *Estrategias y propuestas para la segunda modernización de Andalucía*. Uno de los principales objetivos consiste en fomentar el estudio de lenguas extranjeras para preparar a la población ante los desafíos de la sociedad actual¹⁴.

El Plan de Fomento de Plurilingüismo, aprobado en el año 2005 por la Junta de Andalucía, fue diseñado para promover el aprendizaje de idiomas y la diversidad lingüística en esta región. La búsqueda de un enfoque globalizado en la enseñanza (incluyendo el uso de las tecnologías de la información y de la comunicación, el trabajo cooperativo y las lenguas extranjeras) conlleva la actualización de la formación del profesorado en el plano metodológico-curricular y lingüístico. En este marco se ubica el Máster Universitario en Profesorado en Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas.

El Plan de Fomento de Plurilingüismo pone énfasis en la formación del profesorado tanto de áreas lingüísticas como no lingüísticas¹⁵. También repercute en

¹⁴ «Debemos plantearnos de manera rotunda el reto educativo del conocimiento de otras lenguas en el marco de un mundo global, como instrumento fundamental, para asegurar las posibilidades de éxito y de compartir un futuro mejor» (Consejo Asesor para la Segunda Modernización de Andalucía, 2003: 46).

¹⁵ «Especial mención merece la formación inicial del profesorado de áreas no lingüísticas, en la que habrá que tener en cuenta el nuevo perfil de profesor o profesora que se va a necesitar. En este sentido, es necesaria la inclusión en el plan de estudios de la práctica de las lenguas extranjeras, que podrá considerarse como un importante complemento a la formación de aquel profesorado que desee impartir docencia en los Centros Bilingües. Si además, se le permite desarrollar su período de prácticas docentes en el extranjero (como lo hacen Alemania y Francia), habremos logrado una formación inicial totalmente acorde con este Plan de Fomento del Plurilingüismo» (Consejería de Educación, Junta de Andalucía, 2005: 49).

la formación permanente del profesorado, ya que se tendría que invertir menos esfuerzo, por parte del docente, y menos recursos económicos, por parte de la administración, cuanto mejor sea la formación inicial del profesorado. Algunas de las opciones para la formación permanente son las siguientes: estudios en Escuela Oficial de Idiomas (cursos oficiales presenciales, cursos de actualización lingüística del profesorado y cursos a distancia a través del That's English), cursos de inmersión en verano, licencias de estudio modalidad C, programas de aprendizaje permanente, profesorado acompañante del programa Idiomas y Juventud y ayudas individuales a la formación. En cambio, algunas de esas medidas¹⁶ que se pusieron en marcha, como las licencias por estudio o los cursos de idiomas, se han suprimido por motivos económicos.

El aprendizaje del léxico es fundamental en la enseñanza de idiomas y en los programas de bilingüismo, por lo que se debe fomentar sistemáticamente y desde el nivel más elemental para permitir el intercambio de mensajes cortos desde los primeros días de clase. En los momentos iniciales, de nada sirve la corrección gramatical si el aprendiente no posee los elementos léxicos suficientes para entender los mensajes y para formularlos él mismo, puesto que la información referencial de un texto está contenida en los lexemas. Se podría decir que el léxico transmite el grueso de la información con un esfuerzo relativamente pequeño. Por ejemplo, si un alumno se dirige al profesor diciendo «baño, por favor», o incluso «baño» claramente le estará pidiendo salir para ir al cuarto de baño sin que implique necesidad alguna de formular la pregunta.

Gracias al aprendizaje sistemático del léxico, el estudiante logrará resultados espectaculares referidos con su capacidad de comunicación. En este sentido, numerosos autores abogan por otorgar un papel principal al léxico. Como es el caso de Krashen (1987, *apud* Priyono 2009: 1): «When students travel, they don't carry grammar books, they carry dictionaries» y David Wilkins:

¹⁶ «Acción 48: Gratuidad de los cursos para el profesorado involucrado en el Plan de Fomento del Plurilingüismo en Escuelas Oficiales de Idiomas y otras Instituciones Colaboradoras). [...] Acción 50: Concesión de licencias de estudio específicas (cursos de idiomas en el extranjero, intercambios puesto a puesto, estancias en centros educativos extranjeros, etc.) para la formación en lenguas en el extranjero» (Plan de Fomento del Plurilingüismo, Consejería de Educación, Junta de Andalucía, 2005: 49).

Without grammar little can be conveyed; without vocabulary nothing can be conveyed. If you spend most of your time studying grammar, your English will not improve much. You will see most improvement if you learn more words and expressions. You can say very little with grammar, but you can say almost anything with words! (David Wilkins, 1972: 111):

Para que la enseñanza del léxico sea significativa, es necesario llevar a cabo una recopilación previa de las unidades léxicas que se pretenden enseñar de acuerdo a las características del alumnado. Precisamente, el MCERL (2001: 149) destaca la importancia de la selección léxica para la elaboración de materiales de exámenes y manuales. Pero, ¿cómo se realiza esa selección léxica? ¿cuenta el profesorado con un léxico disponible en lengua española similar en lengua extranjera? ¿qué factores inciden en un mayor o menor caudal léxico?

La búsqueda de la selección del léxico en la enseñanza de idiomas y la evaluación de la competencia léxica del futuro profesorado que va a enseñar esa lengua hacen necesaria la puesta en práctica de una disciplina ya consolidada que responda a las necesidades de esta investigación. En este contexto, se enmarca la elección de la disponibilidad léxica como herramienta necesaria para recabar datos, analizarlos y llegar a conclusiones que se fundamenten en bases teóricas sólidas. Esta disciplina surge en Francia bajo la demanda del Ministerio de Educación Nacional en 1951 de crear un vocabulario de base. Tras la publicación de varios volúmenes, en 1964 aparece una versión mejorada de la obra de Gougenheim, Rivenc, Michéa y Sauvageot, titulada *L'élaboration du français fundamental (1^{er} degré)*.

Esta investigación servirá, no solo para enseñar ese vocabulario de base, sino también para conformar en sí misma la base de un nuevo campo de estudio que cuenta en nuestros días con una vasta literatura. La mayoría de las investigaciones se han realizado en español como lengua materna y tienen como máximo exponente a López Morales quien realiza en 1973 una de las primeras investigaciones sobre el léxico de San Juan de Puerto Rico¹⁷. Además, promueve y dirige el Proyecto Panhispánico de Disponibilidad Léxica (PPHDL) que supone la puesta en común de los aspectos metodológicos en estudios de zonas de habla hispana para la creación de diccionarios que muestren las diferentes sintopías y permitan su contraste.

¹⁷ El primer estudio en español lo realiza Victory (1971) sobre una muestra de estudiantes de 16 a 18 años: 33 monolingües de Monterrey (México), 33 monolingües y 33 bilingües de Houston (Texas).

Grosso modo, este estudio ha seguido la metodología planteada en el Panhispánico con algunas salvedades. Se ha utilizado el sistema de listas abiertas en el que los informantes tienen que escribir todas las palabras que se le vengan a la mente en un tiempo límite de dos minutos por cada centro de interés o estímulo. En cambio, en lugar de utilizar los dieciséis centros de interés clásicos propuestos por Gougenheim *et al.* (1964) se han reducido a nueve para adaptarlos al objetivo de investigación: ocho corresponden a los tradicionales, para poder hacer comparaciones intermuestrales; y uno se incorpora por primera vez en este ámbito por su relación con las nuevas tecnologías¹⁸.

La población objeto de estudio está formada por estudiantes de postgrado, en lugar de estudiantes preuniversitarios, lo que permite analizar la evolución del léxico disponible según el nivel de instrucción al cotejarlo con estudios previos. Asimismo, por primera vez se realiza en España la prueba en español y en lengua extranjera a un grupo de estudiantes que han debido acreditar previamente un nivel B1 según el MCERL. Por ello, se ha ofrecido la posibilidad de realizar la prueba de disponibilidad en el idioma que los hablantes conociesen dando lugar a un léxico disponible en inglés y otro en francés. Asimismo, los informantes han completado un cuestionario con variables socioculturales y una prueba de léxico especializado. De esta forma, se pretende abordar el estudio del léxico disponible desde las tres perspectivas que establece Azurmendi Ayerbe (1983: 187-195): creación de léxicos básicos, estudios de carácter teórico y trabajos de comparación inter e intralingüística.

Esta tesis doctoral focaliza su atención en los factores que condicionan un mayor vocabulario virtual en español y en lengua extranjera de los estudiantes que cursan el Máster Universitario en Profesorado. Interesa conocer el perfil a nivel lingüístico y sociológico de los futuros docentes para indagar en las relaciones que se producen en variables relacionadas con sendos ámbitos. A su vez, trata de evaluar el léxico disponible de uno de los pilares básicos en la implantación del Plan de Fomento de Plurilingüismo: la formación inicial en lengua española y en lengua extranjera del futuro profesorado. Teniendo en cuenta los numerosos aspectos que se

¹⁸ Los centros de interés son: 1) *Partes del cuerpo*; 2) *La ropa*; 3) *Comidas y bebidas*; 4) *La escuela: muebles y materiales*; 5) *La ciudad*; 6) *Medios de transporte*; 7) *Juegos y distracciones*; 8) *Profesiones y oficios*; 9) *Ordenadores e internet*.

pueden abarcar y evaluar en una investigación sobre la formación lingüística del profesorado, en este estudio se fijan los siguientes objetivos específicos:

a) **Ámbito de la sociolingüística:**

- Comparar el léxico disponible con otros estudios para detectar semejanzas y diferencias por razones de lugar, tiempo, nivel de instrucción de los informantes, situación de la lengua (materna, extranjera o segunda lengua), entre otros factores.
- Analizar las variables que condicionan el léxico disponible en español y en lengua extranjera del futuro profesorado para extraer datos que sean de interés en la enseñanza de idiomas.
- Conocer la capacidad de identificación del léxico especializado en español y en lengua extranjera relacionado con los nueve centros de interés del diccionario de disponibilidad léxica y averiguar cuáles son los factores que inciden en un mayor o menor conocimiento del vocabulario especializado.

b) **Ámbito de la psicolingüística:**

- Evaluar las conexiones de palabras que se producen en el lexicón mental del informante a través del programa DispoGrafo¹⁹.
- Realizar un análisis de los procesos de reentrada del léxico al núcleo o prototipo de cada centro de interés mediante el índice de compatibilidad.
- Esbozar una taxonomía de los centros de interés atendiendo a diferentes parámetros.
- Cotejar los análisis anteriores en lengua española y en lengua extranjera para extraer conclusiones sobre el comportamiento del léxico disponible a nivel cuantitativo y cualitativo que comprenda la organización de las palabras.

c) **Ámbito de la lingüística aplicada a la enseñanza de lenguas:**

- Elaborar un diccionario de disponibilidad léxica en lengua española y en lengua extranjera (francés e inglés) sobre nueve centros de interés de gran relevancia en el ámbito educativo.

¹⁹ Programa presentado por Echeverría, Urzúa y Ferreira (2008) que sirve para realizar las relaciones semánticas de corte psicolingüístico.

- Evaluar la corrección ortográfica de los listados de léxico disponible y conocer cuáles son los errores más comunes en los estudiantes para proponer medidas correctoras.
- Seleccionar las unidades léxicas que formarán parte de un vocabulario básico a partir de los listados del léxico disponible.
- Llevar a cabo una propuesta didáctica que se sirva de los resultados de los análisis realizados, sobre todo relacionados con la organización de las palabras en la mente.

Con objeto de dar respuesta a los interrogantes del estudio, se plantean las hipótesis como elemento de conexión entre la teoría y la observación. Asimismo, las hipótesis han guiado la generación de las preguntas del cuestionario sociológico, de la prueba de disponibilidad léxica y de identificación del léxico especializado. De esta forma, es posible corroborar o refutar las mismas por medio de análisis cualitativos y cuantitativos. No se pretende llegar a una constatación unánime y precisa de cada una de las hipótesis ya que la naturaleza de la investigación aporta datos más descriptivos y complejos que un simple *sí* o *no*. En sintonía con los objetivos, las hipótesis se dividen en tres bloques:

a) *Ámbito de la sociolingüística:*

- El promedio de palabras de los preuniversitarios malagueños será inferior al aportado por los estudiantes de postgrado debido a la influencia del nivel de instrucción.
- El caudal léxico en lengua extranjera es inferior en lengua materna (con independencia de la lengua de que se trate, inglés o francés).
- Las estancias en el extranjero, el contacto con los medios de comunicación, la motivación por aprender lenguas y el mayor conocimiento de idiomas inciden positivamente en el caudal léxico.
- Las diferentes variables tendrán una mayor influencia sobre el léxico disponible y la capacidad de identificación de tecnicismos en lengua extranjera que en lengua materna.

- Los resultados de la prueba de disponibilidad léxica y de identificación del léxico especializado guardan una estrecha relación de forma que a mayor caudal léxico mejor será el conocimiento de ese léxico especializado.

b) *Ámbito de la psicolingüística:*

- Las palabras se organizan en redes. La organización de las palabras en subgrupos estará mejor delimitada en español que en lengua extranjera y en los centros más compactos con respecto a los más heterogéneos. Por ello, cuanto mayor sea la cohesión de un centro, mejor definidas estarán las conexiones.
- Los modelos de reentrada del léxico al núcleo varían en función del caudal léxico del informante y del tipo de centro de interés.
- Cada estímulo o centro de interés presenta un diferente comportamiento que marca el índice de cohesión, el número de palabras y el número de vocablos, entre otros aspectos.

c) *Ámbito de la lingüística aplicada a la enseñanza de lenguas:*

- Gracias a la aplicación de la teoría de conjuntos difusos realizada por Ávila Muñoz y Sánchez Sáez (2010, 2011) es posible seleccionar el léxico disponible. El léxico resultante de los seis primeros niveles de compatibilidad léxica corresponde al nivel de usuario básico (A1 y A2) establecido por el MCERL.
- Los estudiantes de postgrado cometen considerablemente menos faltas de ortografía que los preuniversitarios debido al mayor grado de instrucción. La mayoría de las disortografías se deben a una inadecuada acentuación de las palabras.
- Los centros de interés tradicionales solo cubren una parte de las nociones específicas del Plan Curricular del Instituto Cervantes.
- Para elaborar una propuesta didáctica es necesario contar con análisis previos interdisciplinarios.
- El léxico disponible y especializado del futuro profesorado de áreas no lingüísticas es considerablemente menor en lengua española.
- No se exige una preparación previa ni metodológica ni lingüística que forme al futuro profesorado para impartir enseñanza bilingüe.

A través de los resultados se pretende contribuir a las distintas disciplinas que abarca este estudio de disponibilidad léxica. Por una parte, en el ámbito de la lingüística aplicada a la enseñanza de lenguas se dota de una herramienta que permita a los profesionales que se encargan de la planificación didáctica seleccionar el léxico disponible del español²⁰. Por otra parte, se persigue que las conclusiones sirvan de reflexión para mejorar la formación inicial en lenguas extranjeras del futuro profesorado a diferentes niveles: desde la ordenación académica hasta las actividades que puede realizar un estudiante para mejorar su competencia léxica.

En el plano sociolingüístico se han abierto numerosas vías de análisis a través de la inclusión de nuevas variables, tales como la motivación o la realización de estudios en el extranjero que inciden en el léxico virtual en lengua extranjera. Desde el punto de vista psicolingüístico, contribuye con los estudios que analizan la organización de las palabras en la mente concretamente a través de la teoría de las redes semánticas y la teoría de los prototipos. Se propone una taxonomía de los centros de interés que oriente a futuros investigadores a la hora de seleccionar los centros de interés y de realizar comparaciones intermuestrales.

Asimismo, el corpus del léxico disponible supone una fuente de información para futuros estudios dialectológicos y etnolingüísticos que han estado presentes en este estudio de forma transversal. Por un lado, en la edición de los datos ya se comenta la existencia de las unidades léxicas que adquieren un significado diverso en el habla andaluza (*espeto*) o malagueña (*nube*) que podrían ser cotejados con otras sintopías. Por otro lado, en los análisis comparativos generales y sobre las relaciones entre las palabras se han podido conocer los hábitos alimenticios de los jóvenes, su forma de vestir, de pasar el tiempo libre, su relación con las nuevas tecnologías, entre otros aspectos. En síntesis, se persigue mostrar que las diferentes aplicaciones guardan una relación simbiótica y son imprescindibles para realizar una propuesta didáctica.

Tras estas páginas introductorias, en el segundo capítulo se realiza una revisión bibliográfica de los fundamentos teóricos que guían la realización de este

²⁰ En este ámbito, la doctoranda ha realizado varias publicaciones (Santos Díaz, 2013a, 2013b y 2014b) relacionadas con la programación didáctica del español como segunda lengua o como lengua extranjera.

estudio. El primer apartado se centra en la disponibilidad léxica como pilar fundamental que marca la metodología y los análisis de los capítulos posteriores. Los antecedentes comienzan con un análisis descriptivo de la creación del primer léxico disponible de la mano de los autores franceses Gougenheim *et al.* (1964) a través de la consulta de fuentes originales. A continuación se describen los principios epistemológicos y conceptuales de la disponibilidad léxica haciendo hincapié en la evolución de la prueba y en la definición de conceptos clave como *disponibilidad léxica*, *palabra* y *vocablo*.

La encuesta realizada por Gougenheim *et al.* (1964) utiliza el sistema de listas cerradas en la que los hablantes debían aportar un total de 10 palabras relacionadas con un estímulo, denominado en este tipo de estudios *centro de interés*. Dimitrijević (1969) introduce una modificación metodológica por medio de listas abiertas, en las que los informantes podían escribir todas las palabras que se le viniesen a la mente en un tiempo de cinco minutos. La metodología propuesta por el Proyecto Panhispánico incluye el sistema de listas abiertas con una duración de dos minutos para cada uno de los dieciséis centros de interés tradicionales, es decir, aquellos propuestos por los investigadores franceses.

Los parámetros de medición del léxico disponible han evolucionado a lo largo del tiempo así como las herramientas para su cálculo. Este estudio incorpora dos índices novedosos, tan solo presentes en los estudios de Ávila Muñoz y Villena Ponsoda (2010, 2011) que son el índice corregido de descentralización y el índice de compatibilidad léxica. Además, el análisis del vocabulario activo, a través de la prueba de léxico disponible, se ha completado con el análisis del vocabulario pasivo a través de un cuestionario en el que los informantes han debido identificar palabras especializadas con uno de los centros de interés. Para enmarcar la metodología y los análisis efectuados, se han comentado las investigaciones relacionadas con el español como lengua materna y en inglés como lengua extranjera. Los estudios en el mundo hispano se han abordado según el país en el que se realiza la prueba para ilustrar el panorama de los estudios de disponibilidad léxica en las zonas de habla hispana.

A continuación, se han destacado las aplicaciones de la disponibilidad léxica en diversas disciplinas como la etnolingüística, la sociolingüística, la dialectología y la

lingüística aplicada a la enseñanza de lenguas. Precisamente este último ámbito determina el objetivo del primer estudio de disponibilidad y de esta tesis doctoral, por lo que sirve para encadenar el siguiente apartado referido a las variables que inciden en el léxico disponible: la clase social, los hábitos de lectura, el conocimiento de otras lenguas, las estancias en el extranjero, etc. En capítulos posteriores, estos factores serán objeto de análisis cualitativos y cuantitativos para evaluar cuáles son las principales variables que influyen en la capacidad léxica y en la identificación del léxico especializado.

En el tercer capítulo se expone la metodología en la que se enmarca la naturaleza de la investigación, se expone la muestra con detalle y se describen las fases de acceso a los informantes, de recogida de los datos, de edición y de codificación de los datos. En primer lugar, se presenta el cuestionario sociológico sobre el cual se codifican las variables para su posterior análisis estadístico. A continuación, se justifica la elección de los centros de interés en la prueba de léxico disponible y se describe la edición de los materiales estableciendo una distinción entre los criterios comunes al Proyecto Panhispánico (según las recomendaciones de Samper Padilla, 1998) y los criterios de edición propios, tanto generales como específicos de cada centro de interés. Del mismo modo, se explica en qué consiste la prueba de léxico especializado y cómo se han registrado las respuestas de los sujetos. Por último, se presenta en forma de manual de uso el programa DispoGrafo con indicación del tipo de archivo que requiere y las posibilidades que ofrece.

El cuarto capítulo corresponde a los análisis comparativos generales que se dividen en siete apartados. En el primero, se realiza un análisis de los vocablos y palabras aportados en las lenguas del estudio. Partiendo de esos datos, se compara el promedio de palabras por informante entre las pruebas en español, francés e inglés de la muestra de estudio y entre los datos de otras investigaciones. Seguidamente, se presentan los resultados del índice de densidad léxica y el índice de cohesión de cada lengua por separado.

En el plano cualitativo se realiza una comparación intramuestral e intermuestral de los vocablos con mayor índice de disponibilidad según se trate de lengua materna o la lengua extranjera de los hablantes. A continuación, le sigue un

análisis ortográfico donde se extraen los errores más comunes de los informantes y se hace un seguimiento del progreso de los estudiantes preuniversitarios madrileños (Paredes García, 1999) aragoneses (Saura Ramí, 2008) y gallegos (Blanco Domínguez, 2011), pasando por estudiantes universitarios malagueños (Ávila Muñoz, 2007), hasta llegar a los resultados de los estudiantes de postgrado del presente estudio. Por último, se realiza un estudio pormenorizado de los listados del sociolecto femenino y masculino de los centros de interés con diferencias estadísticamente significativas (*La ropa; Comidas y bebidas; La escuela: muebles y materiales*) y que dan paso a los análisis del siguiente capítulo.

En el quinto capítulo, se realiza un análisis estadístico sobre los factores que inciden en el léxico disponible y en la capacidad de identificación del léxico especializado en español y en lengua extranjera. La estructura del capítulo sigue una secuencia lógica que parte de análisis univariantes para presentar las variables más significativas del estudio: variables sociodemográficas y reticulares, variables relacionadas con el aprendizaje de una lengua extranjera y variables lingüísticas. Las dos primeras hacen referencia a las variables independientes que ayudan a explicar el comportamiento de las terceras, es decir, de las variables dependientes. A continuación, los análisis bivariados exponen la relación entre dos variables encaminadas a dar respuesta a las hipótesis de estudio y a seleccionar los factores del análisis multivariante. En este último apartado, se incluyen los análisis de componentes principales, de regresión lineal múltiple y causales que permiten extraer las conclusiones sobre cuáles son las variables con mayor incidencia en la capacidad léxica a través del índice corregido de descentralización léxica (ICoDD) y en el léxico especializado por medio del índice corregido de identificación de tecnicismos (ICIT).

En el sexto capítulo, se analiza la organización de la memoria semántica a través de la teoría de los prototipos y de las redes semánticas. Por un lado, se obtiene el vocabulario prototípico a partir de la selección obtenida mediante la aplicación de la teoría de los conjuntos difusos que, a su vez, se complementa con el estudio de los modelos de reentrada al núcleo. Por otro lado, se analiza la organización de las palabras en subcampos léxicos y las relaciones más frecuentes entre ellas a través del programa DispoGrafo. Por último, se presenta una propuesta didáctica que tiene en

cuenta la selección del corpus lingüístico, los aspectos socioculturales, el nivel de las unidades léxicas según el MCERL y la adecuación de los centros de interés al Plan Curricular del Instituto Cervantes.

El séptimo capítulo recoge las conclusiones finales que sigue la estructura marcada por los objetivos e hipótesis: ámbito de la sociolingüística, ámbito de la psicolingüística y ámbito de la lingüística aplicada a la enseñanza de lenguas. Asimismo, para agilizar la consulta de los objetivos, hipótesis y conclusiones se ha añadido un cuadro resumen en el Anexo II. De esta forma, el lector tendrá acceso a la esencia de esta investigación en solo varias páginas.

Por último, se presentan las referencias bibliográficas y los anexos. En forma de apéndice se ha incluido el cuestionario sociológico realizado, el inventario de las siglas más frecuentes utilizadas, el índice de gráficos y tablas y el léxico nuclear o prototípico de cada centro de interés. Además, se presentan en formato electrónico los diccionarios de léxico disponible ordenados alfabéticamente, según el índice de disponibilidad léxica y el índice de compatibilidad léxica.

CAPÍTULO 2. MARCO TEÓRICO

- 2.1. Introducción**
- 2.2. Orígenes de la disponibilidad léxica**
- 2.3. Principios epistemológicos y conceptuales de la disponibilidad léxica**
- 2.4. Investigaciones relacionadas con el léxico disponible**
 - 2.4.1. En español como lengua materna
 - 2.4.2. En inglés como lengua extranjera
- 2.5. Revisión del modelo matemático**
 - 2.5.1. Cálculo del índice de disponibilidad léxica
 - 2.5.2. Índice de compatibilidad léxica y selección del léxico nuclear
 - 2.5.3. Índice corregido de amplitud léxica
 - 2.5.4. Índice corregido de descentralización léxica
 - 2.5.5. Índice corregido de identificación de tecnicismos
- 2.6. Aplicaciones de la disponibilidad léxica**
- 2.7. Variables que inciden en el léxico disponible**
 - 2.7.1. Sexo
 - 2.7.2. Estancias en el extranjero
 - 2.7.3. Conocimiento de idiomas
 - 2.7.4. Corrección ortográfica
 - 2.7.5. Clase social
 - 2.7.6. Hábitos de lectura y exposición a los medios de comunicación
 - 2.7.7. Contacto y uso de la lengua extranjera en diferentes contextos
- 2.8. Conclusiones parciales**

2.1. Introducción

A lo largo de este capítulo se analiza el estado de la cuestión partiendo de lo general a lo concreto. En primer lugar, se expone la literatura relacionada con la disponibilidad léxica que comienza con un recorrido por las obras pioneras en este ámbito²¹. A continuación, se expone la evolución de la metodología de análisis y de codificación de los datos desde la obra primigenia (Gougenheim *et al.*, 1964) hasta el afianzamiento de las pruebas de listas cerradas con una duración de dos minutos marcado por la proliferación de estudios comunes al Proyecto Panhispánico.

En este sentido, es crucial conocer las investigaciones que se han realizado previamente sobre léxico disponible y, más concretamente, sobre el léxico en español como lengua materna y en inglés y francés como lengua extranjera por tratarse de los idiomas analizados en este estudio. Asimismo, se incluye una revisión del tratamiento matemático de los datos y se definen cuáles son los principales índices utilizados en este estudio y los factores que intervienen en su cálculo.

Una vez comentados los aspectos teóricos y metodológicos, se describen las principales utilidades de los estudios de disponibilidad léxica en diferentes disciplinas haciendo hincapié en la lingüística aplicada en la enseñanza de idiomas. Para terminar, se presentan las variables sociológicas y relacionadas con el aprendizaje y adquisición de una lengua que han sido objeto de estudio con anterioridad: el sexo, las estancias en el extranjero, el conocimiento de idiomas, la clase social, los hábitos de lectura, la exposición a los medios de comunicación y el contacto y el uso de la lengua en diferentes contextos.

2.2. Orígenes de la disponibilidad léxica

En 1947 Aurélien Sauvageot, entonces profesor de l'École Nationale des Langues Orientales Vivantes, acude en representación de Francia a un Comité de lingüistas auspiciado por la Unesco. El objeto de la reunión era promover la

²¹ Como refleja Luque Toro (2006) en el *Léxico español actual*, el campo de estudio del léxico es muy amplio e integra estudios que abarcan neologismos, coloquialismos, léxico arcaico tradicional, léxico innovador, presposiciones, entre otros aspectos.

enseñanza de lenguas para luchar contra el analfabetismo y conseguir difundir una educación de base. En cambio, en aquel contexto la enseñanza de idiomas tradicional en el país galo no establecía discriminación alguna en la selección del vocabulario. Por tanto, a menudo el alumnado se veía expuesto a una cantidad ingente de palabras que no pasaban a formar parte de su repertorio lingüístico.

Con objeto de elaborar un vocabulario de base el Ministerio de Educación Nacional crea en 1951 una comisión y un centro de estudios en l'École Normale de Saint-Cloud, denominado Centre d'Études et de Recherche pour la Diffusion du Français. A la cabeza del centro estaba Gougenheim, profesor de Historia de la Lengua en la Universidad de Estrasburgo quien trabajaba de forma conjunta con Rivenc, entonces profesor d'École Normale. Como fruto de los trabajos científicos de este proyecto el Institut Pédagogique National publica en 1954 un fascículo opúsculo basado en la frecuencia de las palabras (*apud* Gougenheim *et al.*, 1956).

A pesar de que en un primer momento se pensó llamar *Français de base*, las incómodas comparaciones con el *Basic English*²² hacen que se llame *Le français élémentaire*. Cabe mencionar que ya en 1949, Michéa reflexiona sobre la denominación y el contenido de un vocabulario esencial y aboga por sustituir el concepto de vocabulario de base por vocabulario de estructura²³. Aún así, críticas a la obra inglesa, como por ejemplo la realizada por Bongers (1947: 119-134) obligaron a los autores en 1956 a precisar la distinción entre sendos vocabularios y a dedicar un capítulo sobre los orígenes filosóficos del *Basic English*²⁴.

²² El *Basic English*, elaborado por Ogden y Richards en 1928 era un conocido vocabulario que cuenta con un total de 1993 vocablos, de entre los cuales se destaca una lista de vocablo de base con 850 entradas: 600 cosas (sustantivos), 150 cualidades (adjetivos) y 150 operaciones (una mezcla de categorías gramaticales). *Basic* proviene del acrónimo *British American Scientific International Commercial*. Ambos autores cuentan con una vasta obra en la selección del léxico: Ogden, 1930, 1931, 1933, 1935, 1968; Richards, 1935, 1943, 1945, entre otras. Una de las diferencias del *Basic English* con el *Français élémentaire* residía en la metodología de selección del vocabulario: las palabras más comunes en las definiciones de los diccionarios en lugar de las palabras más frecuentes.

²³ «Nous croyons que les éléments d'un voabulaire essentiel doivent être choisis, non seulement pour la part de signification qu'apporte chacun d'eux, mais encore et surtout en considération du soutien qu'ils s'accordent l'un à l'autre dans l'expression cohérente d'un tout physique. C'est pourquoi nous proposerions volontiers de remplacer 'vocabulaire de base' par 'vocabulaire de structure'» (Michéa, 1949: 173).

²⁴ «Une langue semi-artificielle conçue par C.K. Ogden et I.R. Richards dans les années vingt. Il s'agit, en s'appuyant sur les réseaux synonymiques de la langue anglaise, de la réduire aux termes minimaux (premiers) permettant de définir (et donc de comprendre et de s'exprimer dans une certaine mesure) tous les autres termes de la langue, un peu comme dans un dictionnaire. Le Basic est donc un vocabulaire minimal à portée théorique maximale formant un tout, c'est une 'langue universelle à base d'anglais [...] pouvant se suffire à elle-même'» (Gougenheim *et al.*, 1956 : 28).

En 1958 se publica un segundo volumen y en 1959 se reedita *Le français élémentaire* bajo el nombre de *Français Fondamental (1er degré)*, desligándose así de las críticas que lo tachaban de rudimentario, escolar e incluso de peyorativo (Gougenheim *et al.*, 1964: 13). Por último, en 1964 se publica una versión más amplia y mejorada bajo el mismo nombre con algunas variaciones sustanciales.

En el periodo acaecido entre la primera y última obra (1954 y 1964) se trascendieron cambios sociales y políticos que marcaron los objetivos de los vocabularios y las gramáticas de base. En un primer momento, se pretende enseñar francés a los habitantes que forman la Unión Francesa, a los extranjeros y a los inmigrantes que llegaban a Francia. Esa finalidad inicial está vinculada con la creación de la Unión Francesa que nace con la Constitución del 27 de octubre de 1946. El capítulo VIII establece que la Unión Francesa está formada por la Francia metropolitana, los departamentos y territorios de ultramar y los territorios y estados asociados.

Conflictos internos provocaron la separación de algunos de los territorios que formaban parte de la entidad política en vigor: Camboya en septiembre de 1955, Vietnam del Sur en diciembre de 1955, Marruecos en marzo de 1956, Túnez en marzo de 1956 y Laos en mayo de 1957. Además, las guerras de Indochina y Argelia causaron tales estragos que propiciaron la caída de la IV República y el fin de la Unión Francesa en 1958. Por tanto, en la última obra se comenta que a pesar de que uno de los objetivos preliminares era la difusión de la lengua en los países miembros de la Unión Francesa, los propósitos de esa selección del vocabulario pasan a ser los tres que siguen (Gougenheim *et al.*, 1964: 21-22):

- Enseñanza escolar: práctica y cultural.
- Difusión, en los jóvenes estados independientes, de una lengua dada elegida como lengua nacional o lengua extranjera en situación privilegiada.
- Necesidades del viajero, del turista, de la persona que hace prácticas en el extranjero, del extranjero que viene a trabajar a otro país.

En cuanto al método aplicado, el *Français fondamental* reflexiona sobre la metodología estadística llevada a cabo en los dos diccionarios de frecuencia del francés publicados hasta ese entonces en Estados Unidos: *A French Word Book Based*

on a Count of 400.000 running Words de Henmon (1924) y el *French Word Book*²⁵ de Vander Beke (1929). Asimismo, se incluye la referencia a obras derivadas del *French Word Book*, tales como *The French Vocabulary* de Tharp (1939), el *Centrala ordförrädet i franskan* de Schlyter (1951) y el *Basis-Woordenboek voor de Franse* de Verlée (1954). Los diccionarios de Ogden y Richards (1928) y West (1953), combinan las siguientes técnicas (Gougenheim *et al.*, 1964: 13):

- La frecuencia de las palabras en la lengua hablada.
- La búsqueda de las palabras disponibles más útiles.
- El empirismo racional.

El primer aspecto supone un avance metodológico ya que hasta la época todos los vocabularios básicos se basaban en la lengua escrita, en su gran mayoría de carácter literario. En esa misma línea, los autores franceses deciden incluir el registro oral provocando diferencias con otras listas, tales como la lista de Vander Beke²⁶. Para ello, utilizan registros sonoros ya existentes provenientes del Musée de la Parole, Musée des Arts populaires, Archives de la Radiodiffusion y el Musée des Arts et Traditions Populaires. Además, realizan grabaciones directas sobre temas variados que completaron con anotaciones sobre variables sociolingüísticas de los informantes, tales como la profesión, el origen y el nivel educativo.

Las grabaciones dieron lugar a 163 textos con un total de 312 135 palabras diferentes y 7995 vocablos. Previamente, los investigadores contabilizaron las apariciones de los vocablos para incluir en las listas solo aquellos que se hubiesen repetido al menos 20 veces. Una lista se organizó por orden de frecuencia y la otra por orden alfabético incluyendo en ambas la frecuencia y la repartición, es decir, el número de textos en el que el vocablo está presente. Asimismo, las palabras que no aparecen en las listas del *Français élémentaire* por razones justificadas y que cuentan con una frecuencia de al menos 29 apariciones se encuentran marcadas con una cruz. A

²⁵ Disponible en <http://hdl.handle.net/2027/mdp.39015034357353> [Fecha de última consulta el 19 de abril de 2015].

²⁶ Gougenheim *et al.* (1964) analizan las palabras más frecuentes y peculiaridades como las apariciones de palabras sinónimas (*médecin* y *docteur*), de verbos que expresan acciones banales de la vida y palabras gramaticales de carácter literario, entre otras. De forma general, las palabras más cultas tienen una frecuencia superior en la lista elaborada por Vander Beke que en el *Français Fondamental* puesto que el vocabulario ha sido solo extraído de textos escritos.

continuación, se incluyeron dos tablas (Gougenheim *et al.*, 1964: 113 y 115): la primera muestra la relación entre la categoría gramatical y su distribución tanto numérica como porcentual y la segunda analiza la frecuencia de las palabras distribuidas en rangos²⁷.

Las palabras gramaticales aparecen en un mayor porcentaje conforme aumenta la frecuencia mientras que los sustantivos y adjetivos predominan en los vocablos con pocas apariciones. Así, las unidades léxicas con una frecuencia entre 14 000 y 1000 son en su gran mayoría palabras gramaticales: determinantes, pronombres, adverbios, conjunciones y preposiciones que decrecen de 90,9 % en el primer rango hasta 5,7 % en el último. En cambio, más del 50 % de las palabras comprendidas entre 1000 y 20 apariciones son sustantivos. Los adjetivos abundan en el tramo de frecuencia entre 1000 y 60 repeticiones y los verbos presentan una distribución más irregular, según se trate de verbos auxiliares muy usados o verbos más concretos. De hecho, los estudiosos consideran los verbos auxiliares *avoir* y *être* como palabras gramaticales lo que provoca que no exista ningún verbo que sobrepase las 5001 ocurrencias.

Del análisis de los resultados se constata que los sustantivos tienen una frecuencia baja e inestable mientras que las palabras gramaticales y los verbos tienen una frecuencia fuerte y estable. Por tanto, los autores franceses recomiendan incluir en el cálculo de la frecuencia el criterio de la estabilidad de la frecuencia aparte del criterio de la frecuencia absoluta. Asimismo, se cuestionan la metodología de las listas de frecuencias donde la mayor parte de las unidades léxicas corresponden a palabras gramaticales, seguida de verbos, adjetivos y algunos sustantivos de carácter general²⁸.

El hecho que los sustantivos aparezcan de forma reducida en las listas, provoca que palabras que supuestamente son frecuentes presenten un número bajo de apariciones o ni siquiera aparezcan. Para ejemplificar la cuestión, utilizan nombres comunes como *fourchette*, *coude*, *dent*, *métro* y *jupe* cuyo uso está supeditado a una

²⁷ Por ejemplo, el primer rango contabiliza las palabras de más de 5001 apariciones. En ese rango, solo existen 11 palabras, 10 de las cuales son palabras gramaticales más el verbo *avoir*. No aparece ningún adjetivo ni nombre hasta el rango cuarto que comprende los vocablos cuyas apariciones oscilan entre 1000 y 500.

²⁸ «D'une façon générale, les listes de fréquences ne donnent pas ou ne donnent guère les mots concrets. Ceux-ci n'apparaissent qu'avec une faible fréquence. Il y a dans ce fait quelque chose qui, au premier abord, semble surprenant» (Gougenheim *et al.*, 1964: 140).

situación muy concreta de comunicación. Este factor que provoca una baja frecuencia, no conlleva a considerar esas lexías como técnicas o especiales ni a tratar las palabras gramaticales como palabras frecuentes²⁹.

La solución al bajo porcentaje de palabras comunes, pasa por plantear la existencia de dos vocabularios diferentes³⁰. Michéa (1950) establece varias diferencias desde el punto de vista pedagógico entre las denominadas *palabras temáticas* y *palabras atemáticas*. Por un lado, las palabras temáticas, tal y como su nombre indica, están asociadas a un tema o conjunto de temas mientras que las palabras atemáticas aparecen de forma frecuente en cualquier tipo de textos, con independencia del tema tratado. Por otro lado, las palabras temáticas suelen ser sustantivos que sirven para nombrar seres u objetos y las palabras atemáticas se corresponden con palabras gramaticales que ayudan expresar el tema en cuestión. Además, las palabras atemáticas se adquieren por repetición mientras que las palabras temáticas son consideradas un vocabulario de interés que puede ser adquirido por medio de la traducción a la lengua materna o por asociación de ideas (con imágenes, seres u objetos).

Michéa³¹ realizó varias encuestas basándose en el centro de interés, es decir, el tema concreto capaz de activar palabras temáticas. Para la primera prueba, el profesor leyó a 10 de sus 31 alumnos de alemán como lengua extranjera una traducción de un pasaje de «Tonnelier de Nuremberg» de Hoffmann (1987) de dos páginas y media, (alrededor de 800 palabras) y les invitó a dar una lista de veinte palabras que les hubieran llamado la atención. Se constató, que al contrario que en las listas de frecuencia, los sustantivos tanto concretos como comunes eran las palabras más mencionadas mientras que entre las 20 unidades léxicas más mencionadas no se encontraban verbos de carácter general ni palabras gramaticales. Michéa concluyó que en la retención de las palabras los factores con mayor incidencia son el interés y

²⁹ «Ces exemples font apparaître l'ambiguïté de la terminologie banale et non vérifiée qui fait de ces mots des 'mots fréquents' (ce qui est faux) aussi bien que des 'mots usuels', des 'mots courants' (termes commodes, mais qui ne correspondent à aucun concept précis)» (Gougenheim *et al.*, 1964: 141).

³⁰ Anteriormente, Bakonyi (1933) presenta una distinción entre «vocabulario usual» y «vocabulario frecuente» que no convence a los autores franceses.

³¹ *Vid.* Gougenheim *et al.*, 1956.

la atención ya que entre las palabras del estudio aparecían algunas cuyo significado era desconocido por los alumnos.

En otra de las pruebas realizadas, los alumnos tuvieron escribir las 20 palabras que se les vinieran a la mente ante dos situaciones concretas: el momento de espera en la estación para emprender un viaje en tren y las palabras necesarias para escribir una carta. A pesar de que los estudiantes podían escribir cualquier tipo de palabra (palabras gramaticales, adjetivos, sustantivos, verbos y adverbios), entre las 34 lexías más frecuentes solo se encontraban 3 verbos. Los alumnos mostraron cierta preferencia a escribir sustantivos por lo que sustantivos como *voyageur* contaba con 8 menciones mientras que el verbo *voyager* tan solo aparecía en 3 ocasiones. Además, entre las 12 primeras palabras más disponibles de las tres listas, aparecían 7 nombres comunes que se repitían (*gare, train, wagon, paysage, valise, voyageur*).

Una vez más, se corroboró que se produce una inversión de los datos con respecto al vocabulario frecuente: los nombres concretos son numerosos y presentan una estabilidad alta mientras que el número de verbos es reducido e inestable. En cuanto a la dispersión, es decir, la relación entre el número de palabras diferentes y el número total de palabras citadas, se constató que las alumnas de cursos inferiores han extendido las listas a palabras relacionadas indirectamente con el tema. Así, ellas relacionaron el hecho de coger el tren con las vacaciones, las muñecas y otro tipo de asociaciones individuales.

En dos pruebas escolares complementarias, Gougenheim y Michéa seleccionaron el centro de interés de las enfermedades (sin incluir los nombres concretos de las mismas). Michéa realizó esta prueba a un grupo de 21 alumnos de tercer y cuarto curso del instituto de niños de Périgueux mientras que Gougenheim la llevó a cabo con 28 chicas del instituto parisino Hélène-Boucher. Los resultados son muy similares en ambos estudios y corroboran la hipótesis de que la palabra *docteur* no solo es más frecuente sino también más disponible que *médecin*.

Estas encuestas avalan la necesidad de completar el vocabulario frecuente con el vocabulario disponible. Por tanto, Gougenheim *et al.* 1964 recurren al método basado en los centros de interés recomendado por Michéa. En un fascículo publicado

en 1950, Michéa incluye los 16 centros de interés que servirán de base metodológica para elaborar los listados de vocabulario disponibles, siendo los que siguen³²:

- 1) Las partes del cuerpo.
- 2) La ropa (no importa que sea ropa de hombre o de mujer).
- 3) La casa (pero sin los muebles).
- 4) Los muebles de la casa.
- 5) Los alimentos y bebidas de las comidas (en todas las comidas del día).
- 6) Los objetos colocados sobre la mesa y de los que nos servimos en todas las comidas del día.
- 7) La cocina, sus muebles y los utensilios que se encuentran en ella.
- 8) La escuela, sus muebles y su material escolar.
- 9) La calefacción y la iluminación.
- 10) La ciudad.
- 11) El campo o el burgo.
- 12) Los medios de transporte.
- 13) Los trabajos del campo y del jardín.
- 14) Los animales.
- 15) Los juegos y distracciones.
- 16) Los oficios (los diferentes oficios y no los nombres que se refieren a un solo oficio).

Las pruebas se realizaron a escolares de cuatro departamentos diferentes: Dordogne, Marne, Eure, Vendé³³. Para poder realizar un estudio sociolingüístico, se solicitó información sobre el nombre de los sujetos, escuela, sexo, fecha de nacimiento, curso, fecha del ejercicio, profesión de los padres o tutores legales y nivel escolar del alumno (este apartado era aportado por el profesor). Los resultados desvelaron que las diferencias entre departamentos no eran significativas. El factor

³² Traducción literal del francés: 1) *Les parties du corps*; 2) *Les vêtements (peu importe que ce soient des vêtements d'homme ou de femme)*; 3) *La maison (mais pas les meubles)*; 4) *Les meubles de la maison*; 5) *Les aliments et boissons des repas (à tous les repas de la journée)*; 6) *Les objets placés sur la table et dont on se sert à tous les repas de la journée*; 7) *La cuisine, ses meubles et les utensiles qui s'y trouvent*; 8) *L'école, ses meubles et son matériel scolaire*; 9) *Le chauffage et l'éclairage*; 10) *La ville*; 11) *Le village ou le bourg*; 12) *Les moyens de transport*; 13) *Les travaux des champs et du jardinage*; 14) *Les animaux*; 15) *Les jeux et distractions*; 16) *Les métiers (les différents métiers et non pas les noms qui se rapportent à un seul métier)*.

³³ El número total de informantes era de 904 estudiantes (488 chicos y 416 chicas) distribuidos de la siguiente forma: en Dordogne 500, de los cuales la mitad eran chicos y chicas; en Marne 173, siendo 75 chicos y 98 chicas; en Eure 85, compuesto por 73 chicos y 12 chicas; y en Vendée 146, formado por 90 chicos y 56 chicas.

que incidía era sobre todo la procedencia (rural y urbana) y el sexo. Además, se observó bastante similitud entre los resultados en los cuatro listados, sobre todo en el primer centro de interés donde las 20 primeras palabras aparecían casi en el mismo orden.

Siguiendo el criterio de las listas de 20 palabras por centro de interés, se llevaron a cabo varias investigaciones con adultos (carpinteros y lectores de la revista *Vie et langage*). En el primer caso, solo se evaluaron las pruebas sobre los 16 centros de interés anteriormente mencionados de 8 obreros que sirvieron para confirmar los resultados de las pruebas realizadas en escolares. En cuanto a la muestra facilitada por la revista, compuesta por 160 informantes, los lectores debieron aportar verbos relacionados con acciones corrientes que sean usados diariamente. Concretamente escribieron 355 verbos clasificados en tres tipos (Gougenheim *et al.*, 1964: 187):

- a) Verbos que expresan acciones que se llevan a cabo cada día, según un orden determinado.
- b) Verbos que expresan acciones habituales (verbos de movimiento, de percepción, de palabra, de pensamiento, etc.).
- c) Verbos de carácter muy general, susceptibles de ser empleados de maneras muy diversas, tales como *prendre* y *faire*.

Sin embargo, esas pruebas revestirán un carácter complementario ya que las listas del francés básico se basan en las listas de vocabulario frecuente y en los listados de vocabulario disponible aportados por los escolares. Para llevar a cabo la tarea encomendada, el centro de investigación elabora un proyecto con ambos listados que requiere una doble revisión: primero se examinan todas las palabras por orden alfabético y, a continuación, se establece una clasificación de las palabras por tema.

Los criterios de selección del léxico frecuente son detallados con precisión: solo se incluyeron las palabras cuya frecuencia es igual o superior a 29 y que aparecieron en al menos cinco textos³⁴. De esta forma, se evitaba incluir palabras que contasen con una frecuencia elevada pero que apareciesen de forma localizada en

³⁴ A excepción de palabras con frecuencia entre 20 y 29 menciones cuyo interés para la didáctica de la lengua estuviese justificado, tales como *étudiant*, *étranger* y *retard*.

uno o dos textos, como es el caso de *précieux*, *préciosité*, *suceuse* y *suédois*, entre otras. Al aplicar estos dos criterios, el vocabulario se redujo de 1400 palabras a 805 y, posteriormente, se pasó a un segundo filtro compuesto por nueve apartados que eliminó otras 104 palabras más. De forma general, se eliminaron las palabras coloquiales y vulgares, las interjecciones, los sinónimos menos usados, los gentilicios, las palabras no universales, como relacionadas con la religión cristiana y las palabras de un nivel más avanzado.

En cuanto al vocabulario disponible, la selección se realizó ordenando los vocablos según frecuencia de aparición. A pesar de la efectividad en la recolección de los vocablos, los dieciséis centros de interés no dieron respuesta a lo que el centro de investigación y el comité consideraron un vocabulario educativo necesario y universal³⁵. Para paliar las lagunas existentes, la comisión añadió algunos vocablos relacionados con la salud, la enfermedad, la higiene, la limpieza y con términos vinculados con la vida moderna en aquel entonces (*passeport*, *nationalité*, *avenue*, *capitale*, etc.). Además, se eliminaron las palabras consideradas técnicas, como los órganos del cuerpo (a excepción de *cœur*) y vocablos específicos pertenecientes a la vida política, cultural, social y cívica.

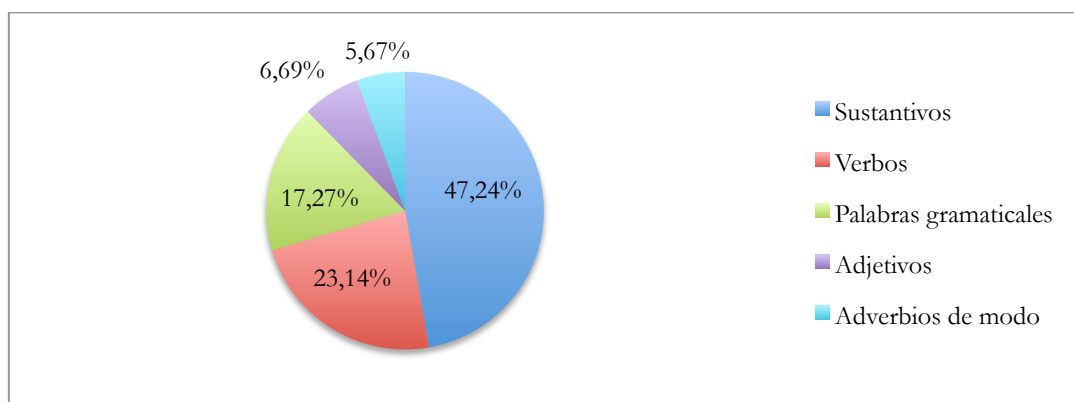


Gráfico 2.1. Distribución del léxico según categorías gramaticales en el estudio de Gougenheim *et al.* (1964)

El resultado de la selección léxica dio lugar a una lista compuesta por 1475 vocablos donde prevalecían los sustantivos. La distribución de las categorías gramaticales es el reflejo de la búsqueda de un vocabulario de conversación práctico

³⁵ «Elle [la comisión] n'a pas voulu que même le 1er degré fût seulement une photographie de la langue usuelle, celle qui sert aux relations sociales les plus simples, au travail habituel, aux achats quotidiens, à l'usage des moyens de transport. Elle a visé plus haut et entendu faire de ce premier degré un premier moyen d'éducation» (Gougenheim *et al.*, 1964: 203).

donde el énfasis estaba puesto en realidades más que en los detalles de las mismas. De ahí que los adjetivos y adverbios de modo apenas supusieran el 13,2 % del total de vocablos.

2.3. Principios epistemológicos y conceptuales de la disponibilidad léxica

Los estudios realizados por Gougenheim, Michéa, Rivenc y Sauvageot para elaborar *Le français fondamental* suponen el nacimiento de una nueva metodología basada en el rango de disponibilidad. El punto de partida de las investigaciones posteriores se centra en el concepto de *vocabulaire disponible* definido por Michéa (cf. 1953: 340)³⁶ como una palabra, que sin ser particularmente frecuente, está siempre disponible y viene a la mente de forma inmediata y natural cuando es necesario.

Esta primera definición ha sido matizada por numerosos investigadores. López Chávez (1994: 69)³⁷ y Samper Padilla, Bellón Fernández y Samper Hernández (2003: 29-30)³⁸ mantienen que ese léxico surge solo en ocasiones concretas y que no es muy frecuente, dejando patente la diferencia establecida entre léxico frecuente y léxico disponible. La interpretación de López Morales (1999a: 11)³⁹ y Bartol

³⁶ «Mais que faut-il entendre par ‘vocabulaire disponible’? Un mot disponible est un mot qui, sans être particulièrement fréquent, est cependant toujours prêt à être employé et se présente immédiatement et naturellement à l’esprit au moment où l’on en a besoin. C’est un mot qui, faisant partie d’associations d’idées usuelles, existe en puissance chez le sujet parlant, dès que ces associations entrent en jeu. L’existence d’un vocabulaire disponible n’est pas une vue de l’esprit. Il suffit de s’observer soi-même lorsqu’on parle et surtout lorsqu’on écrit pour se rendre compte que, souvent, plusieurs termes entrent en concurrence au moment de l’expression. Un seul est employé. Les autres ont simplement marqué discrètement leur présence» (Michéa, 1953: 340).

³⁷ «El conjunto de vocablos que utiliza una comunidad de hablantes como parte medular de su léxico patrimonial pero que únicamente aparecen en situaciones precisas, por lo que no son muy frecuentes pero sí muy conocidos. Son inestables debido a que es necesario que se toque un tema específico en el discurso para que aparezcan» (López Chávez, 1994: 69).

³⁸ «El vocablo disponible se caracteriza porque se presenta en la mente del hablante de forma inmediata y natural cuando se trata de un tema determinado. Es una palabra que, sin ser necesariamente frecuente, vive potencialmente en el hablante y se actualiza en cuanto se producen ciertas asociaciones» (Samper *et al.*, 2003: 29-30).

³⁹ «Existe en el lexicon mental una serie de términos que no se actualizan a menos que sean necesarios para comunicar una información muy específica. Se trata de un léxico disponible, cuyo estudio no puede emprenderse manejando frecuencias, porque este factor es pertinente sólo en el caso de las actualizaciones léxicas efectivas, no de las potenciales» (López Morales, 1999a: 11).

Hernández (2010: 91)⁴⁰ incluye el concepto de lexicón mental entendido, *grosso modo*, como el lugar donde se almacenan las palabras. Más allá de la creación de un diccionario de léxico disponible, el análisis de las asociaciones entre el estímulo y las palabras producidas por los informantes es de gran relevancia para conocer la estructura del léxico mental.

Para activar ese léxico disponible en situaciones concretas, tanto los autores franceses como sus sucedáneos se basan en los denominados centros de interés. La bibliografía de referencia en torno a los centros de interés puede consultarse desde la perspectiva lexicográfica y pedagógica. Desde el punto de vista de la lexicografía, existen obras medievales que recogen la organización de los vocablos en bloques temáticos denominados *Vocabularia*. El objetivo era facilitar la adquisición del latín en un momento donde la gramática y el léxico jugaban un papel predominante en la enseñanza de idiomas.

En el ámbito educativo, su precursor es Decroly (1871-1932) quien organiza los diferentes contenidos curriculares en torno a esos centros o temas concretos según la edad del alumno. Ejemplo de esa metodología, se encuentra en el libro de la pedagoga argentina Guillén de Rezzano (1958) donde desarrolla los centros de interés correspondientes a los grados infantiles superiores haciendo referencia al nivel académico del alumnado «niños que ya saben leer y escribir y de una edad que oscila entre los siete y ocho años» (Guillén de Rezzano, 1958: 12). Para el primer grado (siete años) desarrolla aspectos relacionados con la alimentación de forma elemental: *la granja; el establo, la vaca, la leche; la conejera; el chiquero; la colmena; el gallinero; la huerta; el granero, el trigo, el pan; la primavera*⁴¹. Para el segundo grado, incluye los centros que se relacionan con la vivienda incluyendo la geografía: *la vivienda; la montaña; el río; la llanura; el mar; la selva, el árbol; calefacción y alumbrado; viajes: medios de transporte*.

Ambas disciplinas seleccionan y delimitan determinados aspectos de la realidad con fines pedagógicos encaminados al diseño curricular en general o a la

⁴⁰ «El léxico disponible supone un nuevo enfoque sobre el léxico. No se trata de las palabras utilizadas al producir determinados textos; sino de las palabras que, organizadas, están en nuestro lexicón mental y están ‘disponibles’ para ser utilizadas cuando las circunstancias comunicativas lo requieran; cuando un estímulo active el nodo cerebral adecuado» (Bartol Hernández, 2010: 91).

⁴¹ En este estudio se ha optado por incluir los nombres de los centros de interés en cursiva siguiendo el estilo marcado por la obra pionera Gougenheim *et al.* (1964) secundada por López Morales (1995), entre otros.

selección del vocabulario para el aprendizaje de una lengua. En el caso de la disponibilidad léxica, Michéa incorpora la definición del concepto de centro de interés haciendo alusión a la pedagogía tradicional. De hecho, se observa cierta similitud entre algunos de los centros presentes en obras como la de Guillén de Rezzano (1958) y los centros propuestos por el profesor francés: *la vivienda, la huerta, calefacción y alumbrado, medios de transporte y el pan*, entre otros. Si bien es cierto que el concepto podría hacer referencia a obras lexicográficas, el carácter ideológico del *Français Fondamental*, que conlleva la alfabetización y aculturación a través de la enseñanza del francés, toma como punto de partida estudios pedagógicos (Michéa, 1953: 39)⁴².

Los centros de la obra de Michéa (1953) tomados como referencia para los estudios del *Français Fondamental* son dieciséis. Sin embargo, cada investigador mantiene, disminuye o aumenta el número y tipo de centros en función de los objetivos de los estudios. La primera investigación que incluye la metodología francesa es llevada a cabo en 1964 por Pfeffer, especialista en lexicometría alemana. El estudio tiene un objetivo meramente didáctico: aprendizaje del alemán hablado contemporáneo. No obstante, la disponibilidad léxica se constituye como uno de los cuatro criterios de selección del vocabulario gracias al aporte de nombres concretos útiles para el aprendizaje y la enseñanza de idiomas.

La encuesta se realiza en la principal zona germanófona de Europa Central y comprende 82 escuelas en 48 ciudades diferentes de Alemania, Austria y Suiza. La edad de los alumnos oscila entre los quince y dieciséis años y la proporción por sexos es similar, elevándose a casi 2000 participantes. Para la recogida del léxico disponible, se solicita a los informantes que escriban en dos pruebas de diez minutos una lista de veinte sustantivos y una lista de doce verbos y ocho adjetivos relacionados con cada centro de interés. Esos cuestionarios recogen en total 833 000 palabras y 34 000

⁴² «L'exploration statistique des centres d'intérêt est devenue une nécessité. À la vérité, l'expression 'centre d'intérêt' désigne, dans la pédagogie traditionnelle, des groupements de forme variée. Tantôt il s'agit d'une même classe d'objets (parties du corps, vêtements, animaux, plantes), tantôt plus proprement de 'milieux' (la ville, le village). Ces ensembles offrent un choix plus ou moins large ; ils permettent à la pensée de s'engager dans un plus ou moins grand nombre de directions. En un mot, ils peuvent avoir une 'cohérence sémantique' différente. Il est donc normal que les probabilités de rappel ne se hiérarchisent pas exactement de la même façon dans les différents centres» (Michéa, 1953: 39).

vocablos, entre los que se excluyen los nombres propios, topónimos y adjetivos derivados de nombres de lugares.

El vocabulario básico alemán está organizado por orden alfabético, según partes del discurso, frecuencia y origen. Los materiales usados han sido 400 grabaciones de entrevistas y 5400 encuestas escritas de asociación espontánea. Las palabras de la prueba de disponibilidad léxica han sido seleccionadas según su amplitud de repartición y su índice de disponibilidad (superior a 100 menciones) sumando un total de 347 de las 1269 palabras que conforman ese vocabulario básico.

Ese carácter complementario que aportan las pruebas de disponibilidad léxica, se ve reflejado en investigaciones posteriores como la irlandesa, *Buntus Gaeilge*, publicada en 1966⁴³. La obra está escrita en irlandés y su título hace referencia a la base y principio del ese idioma «a nucleus of Irish» (Ó Huallacháin, 1978: 509). El director del trabajo, Ó Huallacháin (1978: 502), afirma que en el régimen británico la enseñanza del irlandés en las escuelas ha sido autorizada en 1878 y la enseñanza en lengua irlandesa en 1904, de ahí la necesidad de recopilar material necesario dirigido a profesores y organizadores de cursos.

El trabajo se divide en tres partes: el estudio de las estructuras sintácticas, el vocabulario y morfología y las palabras disponibles. Se centra en el irlandés hablado por lo que para la recopilación de las palabras y la grabación de las conversaciones se basa en las técnicas empleadas en el *Français Fondamental*.

Hundreds of thousands of words of recorded conversations from Irish-speaking areas provided the basis for the study of vocabulary, verb-morphology and syntax which was designed to show the kind of Irish which should be taught in order that an initial measure of performance would be attained as soon as possible. The techniques employed are an elaboration of those pioneered by the CREDIF team which produced *Le Français Fondamental* (Ó Huallacháin, 1978: 9-10)⁴⁴.

Concretamente, la prueba de léxico disponible la realizan 260 irlandeses y 260 ingleses de Irlanda sobre 20 centros de interés⁴⁵ que anotaron más de 200 000

⁴³ Los datos del estudio se han recabado a partir de la investigación de Njock (1978) y las publicadas por Ó Huallacháin en 1970 y 1978.

⁴⁴ El número de página corresponde a una anotación manuscrita. El documento ha sido extraído de la base de datos especializada en educación del *Education Resources Information Center*: <http://eric.ed.gov/?id=ED041256> [Última consulta el 10 de abril de 2015].

⁴⁵ 1) *The body and its parts*; 2) *physic exercises, athletics, sports*; 3) *recreation hobbies, indoor pastimes*; 4) *illnesses (kinds of sicknesses), injuries, cures*; 5) *religion, the church*; 6) *clothes of all kinds (men, women, boys, girls)*; 7) *food and drink*; 8) *meals, utensils for each meal and for preparing meals*; 9) *the home or house: furniture, fittings, heat, light*; 10) *business: buying, selling*;

palabras. Cada sujeto tenía que escribir las 20 palabras que considerara más útiles (nombres, verbos o adjetivos) para cada uno de los centros de interés. Los informantes fueron seleccionados teniendo en cuenta seis factores: la edad, el sexo, la residencia (zona rural o urbana), la ocupación y la distribución geográfica. La distribución de los hablantes por niveles educativos es la siguiente: 100 provienen de educación primaria, 60 de escuelas comerciales (*trade schools*), 50 de enseñanza secundaria y 50 adultos de los cuales 25 son estudiantes. El factor del reparto geográfico encuentra su importancia en la distribución equilibrada de los sujetos irlandeses según su dialecto. Por ejemplo, para los irlandeses que hablan irlandés gaélico no tenía sentido tener en cuenta el factor *urbano/rural* puesto que la casi totalidad de las personas forman parte de comunas rurales.

En 1969 el yugoslavo Dimitrijević publica la primera obra de referencia sobre disponibilidad léxica en inglés bajo el nombre de *Lexical Availability. A New Aspect of the Lexical Availability of Secondary School Children*. La muestra del estudio estaba formada por 185 alumnos de 14 años⁴⁶ (89 chicos y 96 chicas) de *Junior Secondary Schools* (125, de los cuales 59 chicos y 66 chicas) y *Senior Secondary Schools* (30 chicos y 30 chicas) de zonas rurales y urbanas de Edimburgo. El número de centros se redujo a once, siendo los que siguen:

- 1) Animals.
- 2) Countryside.
- 3) Town.
- 4) Entertainment.
- 5) Jobs and professions.
- 6) Science.
- 7) Means of transport.
- 8) Politics.
- 9) Parts of the house.

11) *life in town and city*; 12) *life and work in the country and the landscape*; 13) *animals, birds, fish, insects*; 14) *trees, plants, flowers, weeds, crops*; 15) *trades and kinds of occupation or work*; 16) *travel, means of transport and things connected with them*; 17) *the family, relatives and people's feelings about themselves and others*; 18) *school life, activities, furniture*; 19) *the ruling of the country, government, politics*; 20) *the weather, the time of day, seasons*.

⁴⁶ Todos los niños tienen 14 años a excepción de 3 chicas y 2 chicos que tienen 15, de ahí la generalización por parte del autor.

10) Food and drink.

11) Clothes.

La elección de cada centro de interés atiende a criterios debidamente justificados por el investigador en cada caso haciendo hincapié en la familiaridad del tema y su representatividad⁴⁷. Se han mantenido 10 centros de interés tradicionales, es decir, que aparecen en la obra pionera y se ha añadido un nuevo centro, correspondiente al número ocho, *Politics*. El motivo de la eliminación de ciertos centros de interés atiende a cambios introducidos en la metodología y al límite temporal permitido para hacer las pruebas, concretamente dos clases o sesiones. Una de las variaciones consiste en establecer un tiempo límite en cada centro de interés para que el hablante pueda escribir todas las palabras que conozca de forma que el total de unidades léxicas pudiera servir de indicador de la fluidez verbal (*cf.* Dimitrijević, 1969: 5)⁴⁸.

El autor marca la diferencia entre las palabras disponibles y las parcialmente disponibles según su grado de disponibilidad (*cf.* Dimitrijević, 1969: 4)⁴⁹. Por un lado, las palabras disponibles son más familiares y no requieren trabajo localizarlas bien porque el hablante no tenga que pensar en buscar la palabra cuando la quiere utilizar en el discurso o porque no siente el esfuerzo producido al buscarla. Por otro lado, las palabras parcialmente disponibles son aquellas el individuo conoce y utiliza pero no siempre están disponibles. No son fácilmente recordables y requieren un esfuerzo consciente encontrarlas por lo que no todos los informantes de las pruebas las incluyen entre sus respuestas.

Para determinar el tiempo necesario, Dimitrijević realiza un estudio piloto en Queensferry donde asigna cuatro minutos a cada centro de interés. Comprueba que

⁴⁷ «The centre of interest had to deal with those words that are familiar to children, that can be arranged without much overlapping between the centres and which together represent a considerable part of children's vocabulary» (Dimitrijević, 1969: 16).

⁴⁸ «The test had a time limit, each centre of interest being allotted the same length of time. The informants were asked to give «al the words they knew, as fast as they could». This means, in other words, that the total output of lexical items in one centre, by one pupil, could be considered as a kind of indicator of his/her verbal fluency» (Dimitrijević, 1969: 5).

⁴⁹ «This distinction between available and unavailable words brings us to the notion of the degree of availability. A word has 100 % degree availability if it is available to all informants who take part in a test. The fewer the number of informants who give a word the smaller degree of availability of that word for that population, or the easier a lexical item is recalled by a person, a group of people, the more available this lexical item is (the higher degree of availability it has) and v. v.» (Dimitrijević, 1969: 4).

solo algunos alumnos agotaban el tiempo permitido sin dificultad mientras que la mayoría terminaba antes o retomaba la tarea tras una pausa. Como la población de estudio pertenecía a un medio rural y, según el autor, su nivel académico era más bajo que el de los alumnos de Edimburgo, decide aumentar en un minuto la prueba. De esta forma, no solo se conseguiría obtener realmente el léxico disponible del alumnado sino también la capacidad de producir las palabras⁵⁰.

Del primer ensayo se desprende la importancia conferida a los factores que influyen en la disponibilidad léxica. Las variables incorporadas son el sexo, el medio (rural y urbano), el tipo de centro educativo (distinguiendo entre *Junior Secondary Schools* y *Senior Secondary Shools*), el estrato social (dividido en cinco grupos aunque ningún informante perteneciese al primero)⁵¹ y los resultados C. Q. (que combina el cociente de inteligencia IQ que realizan los estudiantes antes de empezar la secundaria junto con la evaluación de los docentes). No obstante, el análisis cualitativo y cuantitativo del artículo solo recoge el sexo y el estrato social de los alumnos.

Para el recuento de las palabras, Dimitrijević introduce el concepto de «lexical item», término que precisa de la distinción entre *vocablo* y *palabra*. En el lenguaje no especializado existe la tendencia a usar ambos términos como sinónimos. Buen ejemplo de ello se encuentra en la primera acepción de vocablo en el *Diccionario de la lengua española* puesto que aparece el término palabra en forma de hiperenlace⁵². Por tanto, no solo se indica que son sinónimos sino que además redirige al significado de palabra para definir vocablo. Aunque etimológicamente estos términos provienen de voces diferentes, ambas se usan como equivalentes desde el origen de la lengua, como muestra el Diccionario crítico etimológico de Corominas.

⁵⁰ «The aim of this investigation was to find out what words were most available to boys and girls, and how many such words. If we wanted to elicit such words, i.e. not those words which the pupils knew and did not use, but only those words that were always available to them, under any circumstances, the test had to be limited in time» (Dimitrijević, 1969: 14).

⁵¹ La obra de referencia es la categorización se titula *Classification of occupations*, 1960, General Register Office, London, Her Majesty's Stationery Office.

⁵² La consulta se ha realizado en línea el 20 de abril de 2015 en la versión electrónica del *Diccionario de la lengua española*.

PALABRA, del antiguo *parabla* y éste del lat. *parabŏla* ‘comparación, símil’, que a su vez vino del gr. *Παραβολή* ‘comparación, alegoría’ (de *παράδᾶλλειν* ‘poner al lado, comparar’; *βάλλειν* ‘echar’); en romance se pasó de ‘comparación’ a ‘frase’ y luego ‘vocablo’. 1.^a doc.: orígenes del idioma (Cid, etc.).

De uso general en todas las épocas y común a todos los romances de Occidente. La forma arcaica *paravla* o *parabla* se lee todavía en Alex. El sentido etimológico se nota todavía en textos arcaicos: en Berceo Mil, 16b es ‘frase’ o ‘comparación’, en Alex., 401 será ‘refrán’ o ‘sentencia’; lo más común es hallar desde el principio (p. Ej. En el Cid) usos semánticos que no difieren esencialmente de los modernos, aunque también en ellos se podría traducir palabra por ‘frase’ (estas palabras fabló) (Corominas, 1974 [1954], tomo IV, p. 616).

Por su parte, la entrada de vocablo se inserta como un derivado culto de *vŏcāre* ‘llamar’, que proviene de la raíz latina *vŏx*.

VOZ, del lat. *vŏx*, *vŏcis*, ‘sonido producido por el aire expelido por los pulmones al hacer vibrar las cuerdas vocales’. 1.^a doc: Cid.

Vocal [letra vocal, hacia 1250, Setenario, 4.20; A. DE PAL. 535B: «vocales se dizen las letras que de por sí fazen boz»; Nebr.; oración vocal, Sta. Teresa, los vocales del Concilio, fin S. XVI, Aut.], tomado de *vŏcālis* ‘hecho con la voz’, ‘sonoro’, ‘vocal’; vocálico, vocalismo; vocalista ‘el cantor, músico vocal’ [1834]; vocalizar; vocalización; vocalizador.[...]

Derivados cultos de *vŏcāre* ‘llamar’ (de la misma raíz que *vŏx*): Vocablo [APal. 9d, 141d, 535b; Nbr.; en escrituras en bajo latín encontramos vocabulo (o vogabulo 1081, Oelsch.) con el sentido ‘por nombre, llamado’], de vocabulŭm ‘denominación’, ‘nombre’, ‘palabra’, ‘término’; vocabulario [«v. Vocabularium» Nebr.] (Corominas, 1974 [1954], tomo V, p. 845).

En el ámbito de la lingüística estadística, en el que se inserta la presente investigación, se incluye la distinción entre *vocablo* y *palabra*. Según Müller (1969: 31), resulta muy complejo delimitar el segmento del discurso y máxime cuando se trata de los términos *vocablo*, *palabra* y *unidad léxica*⁵³. Consciente de que cada autor establece unas normas similares para diferenciar las unidades léxicas, Müller (cf. 1969: 32)⁵⁴ aboga por la existencia de una norma convencional como la del sistema métrico

⁵³ «Inutile d'insister sur les incertitudes de la linguistique en ce qui concerne ce 'segment de discours' que le sens commun identifie comme étant 'un mot'. Dire d'une phrase qu'elle est formée de 10 mots, ou dire que le texte du Cid, par exemple, en compte 16 690, cela suppose soit une définition assurée du mot comme unité de texte, telle que jamais celui qui compte ne puisse hésiter ou ne doive arbitrer de façon subjective, soit une norme conventionnelle prévoyant tous les cas douteux. [...] Les difficultés ne font que croître quand on passe de la délimitation du mot à celle du vocable, unité de lexique» (Müller, 1969: 31).

⁵⁴ «On ne saurait épuiser cette question 1. Disons seulement qu'à l'heure actuelle, en français comme dans d'autres idiomes, chaque auteur de dépouillements lexicaux est contraint de se poser la question et, finalement, de construire sa propre norme. [...] On ne peut que souhaiter, au moins pour le français moderne, l'adoption d'une norme conventionnelle, même discutable aux yeux du linguiste, mais acceptée comme le fut le système métrique. En attendant, il est nécessaire que chaque chercheur indique, au moins sommairement, sur quelles bases il a travaillé» (Müller, 1969: 32).

sobre la cual pudiesen basarse los estudiosos. Mientras tanto, propone al investigador la elección de unos determinados criterios debidamente argumentados sobre los que fundamentar su estudio.

Müller (1963: 31-32) va más allá y distingue entre dos acepciones del término genérico *palabra*. Por una parte, hace mención a la unidad que constituye el texto y, por otra parte, a la unidad de léxico. A la primera acepción corresponde literalmente al término *palabra* y, a la segunda se ajusta *vocablo*, entendido como los términos que aparecen en la nomenclatura de un diccionario y que forman el vocabulario de una lengua. Trasladando ese enfoque al estudio de disponibilidad léxica, *los vocablos* se refieren al número de palabras diferentes y *las palabras* a las unidades de texto producidas por los informantes, es decir, al número de ocurrencias de los vocablos. En la mayoría de las investigaciones de disponibilidad léxica se añade esta distinción, de manera más o menos explícita, coincidiendo con las siguientes definiciones:

Palabra [...] cada una de las unidades léxicas empleadas, tanto aisladas como unidades complejas o combinación de lexemas para designar un referente, repetidas o no, que integran un texto (Ayora Esteban, 2006: 87).

Vocablo[...] cada una de las unidades léxicas distintas que registra un diccionario de frecuencias. Proporciona información sobre las categorías gramaticales con que se reflejan en el acto discursivo (Ayora Esteban, 2006: 87).

A pesar del consenso a la hora de utilizar *vocablo* y *palabra* en los estudios de disponibilidad léxica, en algunas investigaciones se realizan puntualizaciones sobre su uso y que se completa con el término *unidad léxica*. Según Dimitrijević (1969: 4-5)⁵⁵, en ocasiones las palabras contabilizadas en los estudios de disponibilidad léxica y entendidas como unidad de texto, comprendían más de una palabra. Por tanto, el término más adecuado sería el de *unidad léxica*, como unidad de significado que puede estar formada por más de una palabra. Ejemplo de ello, son las entradas léxicas formadas por más de una palabra, tales como, *intestino grueso*, *disco duro*, *ama de casa*, etc.

⁵⁵ «We have just said that there was no restriction in the test, as to what words had to be given, as far as they were available to the informants. This had led to the elicitation of one, two or more words which had to be taken as one single 'unit' or response, such as 'railway station', 'fresh air', etc. Because of such 'multi word responses' we have adopted the term 'lexical item' denoting all kinds of responses- those responses consisting of one, two or more words. Other alternative terms could be 'lexical phrase', 'lexical term' o 'lexical unit'» (Dimitrijević, 1969: 4-5).

En esta misma línea, Gómez Devís (2004: 43), analiza el significado de unidad léxica recogiendo la perspectiva de diferentes investigadores y concluye que en el ámbito de la estadística léxica el vocablo se entiende como unidad léxica.

Una vez esbozados estos conceptos clave, a modo de inciso, a continuación se reanuda la presentación por orden cronológico de las investigaciones que marcan de una forma u otra la metodología de la disponibilidad léxica actual. Este es el caso del trabajo publicado en dos tomos por Mackey (1971a): el primero titulado *Le vocabulaire disponible du français. Le vocabulaire concret usuel des français et acadiens: étude témoin* y, el segundo, *Le Vocabulaire disponible des enfants acadiens*. Se podría considerar como la primera investigación basada en la obra pionera puesto que las indagaciones iniciales remontan a los años 1958-1959⁵⁶.

El estudio se inicia a través de un proyecto en las regiones de Quebec y Manitoba respaldado por el Centre International de Recherches sur le Bilinguisme de la Université Laval situado en Quebec con objeto de comparar las variedades léxicas de diversas regiones francesas y canadienses. En algunas regiones bilingües, los hablantes realizan la prueba con cinco centros de interés complementarios en inglés. Debido a la falta de consenso metodológico entre las universidades implicadas, se crea un comité interuniversitario que en 1963 marca las pautas de recogida, procesamiento y análisis de datos y consensúa las divergencias. Por ejemplo, los centros de interés seleccionados en Montreal y en Laval eran diferentes ya que en algunos casos se unificaban algunos centros tradicionales y en otros se añadían nuevos. Para permitir el parangón con los datos recopilados en *Le Français Fondamental* y a su vez dar respuesta a los objetivos del proyecto canadiense, se mantienen los dieciséis centros tradicionales y se añaden los siguientes: 17) *l'église et la vie paroissiale*, 18) *les sports et l'équipement*, 19) *les outils*, 20) *l'automobile et ses parties*, 21) *actions courantes* y 22) *les affaires*.

Es precisamente la fidelidad a la metodología aplicada en la investigación francesa la que marca el inicio de las comparaciones intermuestrales en el campo de

⁵⁶ El proyecto de investigación de Dimitrijević (1969) fue iniciado con posterioridad, en 1963, aunque la publicación fue anterior a la de Mackey (1971).

la disponibilidad léxica a través del léxico disponible de Francia y de Canadá⁵⁷. Sin embargo, en el prefacio el autor expone que más que presentar una comparación entre dos vocabularios en sí misma, la obra intenta explicar la forma de llevarla cabo para proporcionar las herramientas adecuadas a futuros investigadores⁵⁸.

Por una parte, se compara el grado de disponibilidad de las palabras en las dos listas y, por otra parte, se analizan las palabras que solo están presentes en uno de los vocabularios. Para acometer tal objetivo, asistentes de lingüística y estudiantes corrigieron a mano las listas impresas por orden alfabético tanto del vocabulario de Francia como de Acadia para evitar fundamentalmente tres tipos de situaciones:

- Una palabra se presenta en singular en una lista y en otra en plural.
- Una palabra aparece en género masculino y en otra lista en femenino.
- Una palabra compuesta, una locución o un grupo de palabras pueden ser clasificadas como palabras únicas en una lista y en la otra no. Esas palabras compuestas pueden aparecer con o sin guión, por ejemplo el caso de *pois* y *petit pois*, *football* o *foot-ball*.

Surgen así las primeras reglas o pautas de lematización para permitir que dos vocablos iguales, como *pied* y *pieds* no fuesen considerados como lexías diferentes. Se fija la homogeneización de las palabras en número singular y género masculino pero en el caso de las palabras compuestas no se establece una regla fija.

Las encuestas fueron realizadas mediante listas cerradas de 20 palabras a niños de entre 9 y 12 años (la misma edad de los estudios franceses) de 20 clases diferentes en 12 escuelas. Cabe destacar que en un primer momento se recopilaban datos de 1745 alumnos de 9 a 18 años, repartidos en 47 clases y 19 escuelas en las diversas regiones francófonas del Nouveau-Brunswick pero se eliminaron los resultados de los niños de 13 a 18 años para facilitar la comparación con el vocabulario de Francia. Asimismo, en algunas regiones bilingües, los sujetos realizaron la prueba de léxico

⁵⁷ «Il a pour but non seulement de fournir la documentation pour une comparaison du vocabulaire commun de deux régions francophones, mais également de servir d'étude témoin à d'autres études analogues» (Mackey, 1971a: 13).

⁵⁸ «L'étude ne va pas assez loin pour établir les comparaisons entre les deux vocabulaires. Elle se borne à fournir les outils à ceux qui entreprennent une telle comparaison. Elle n'exploite pas la documentation; elle explique simplement comment cette documentation a été obtenue» (Mackey, 1971a: 11).

disponible con cinco centros de interés complementarios en inglés. De forma general, cada estudiante realizó la prueba sobre 22 centros, cada uno en una hoja diferente, en un tiempo máximo de 15 minutos por centro. Además, proporcionaron información de interés para la sociolingüística relativa al sexo, la edad, las lecturas que realiza y el medio donde viven (rural o urbano), entre otras variables.

En cuanto a las comparaciones intermuestrales, se centran en la edad, la influencia del bilingüismo y sobre todo en las diferencias geográficas, para lo cual se divide el país en doce regiones: Gaspésie, Bas-du-Fleuve, Lago Saint-Jean, Quebec Noroeste (Abiti, Hull, etc.), región de Montreal, Richelieu, Sherbrooke, Trois-Rivières, Quebec, Ontario, Oeste de Canadá y Acadia, siendo esta última donde se centra el objeto de estudio⁵⁹. Se prueba la existencia de peculiaridades culturales y sociales en la disponibilidad léxica como fruto de la influencia del medio donde vive el hablante. De ahí la importancia de esta disciplina no solo en el ámbito de la enseñanza de idiomas sino también de la psicología y sociología.

En ese mismo año, en 1971⁶⁰ se publica la tesis de Victory (Jr.) titulada *A study of Lexical Availability among Monolingual-Bilingual Speakers of Spanish and English*. La metodología del estudio se basa en la investigación de Dimitrijević (1969) aunque con ligeros cambios, como la realización de las pruebas en dos lenguas⁶¹. La población analizada está formada por 99 estudiantes de 16 a 18 años, 33 monolingües en español (de Monterrey, México), 33 monolingües en inglés (de Houston, Texas) y 33 bilingües (también de Houston). Además de esa variable, los hablantes aportan información sobre el nombre, la edad, el sexo, la lengua o lenguas habladas en casa y la clasificación ocupacional del padre o de la persona que sustenta la familia.

El principal objetivo de la investigación consiste en encontrar las diferencias mostradas a nivel cuantitativo y cualitativo en las palabras disponibles aportadas por los tres grupos en diez parcelas diferentes del vocabulario activo. De forma complementaria, se busca la aplicación práctica de estos estudios léxicos que

⁵⁹ Recopilaron datos de 2000 hablantes de 9 a 18 años pero solo seleccionaron las encuestas de aquellos entre 9 y 12 años.

⁶⁰ La investigación se puso en marcha en 1970 de ahí que tome como referencia el estudio publicado solo un año antes por Dimitrijević (1969).

⁶¹ «The technique used follows closely behind those employed by a 1963 investigation of lexical availability performed in two Edinburgh, Scotland schools by the Yugoslav educator, Dr. Naum. R. Dimitrijević, of Belgrade» (Bailey Victory, 1971: 5).

considera clave en varios ámbitos: para la creación de diccionarios y manuales, para la elaboración de pruebas lingüísticas y para la selección del vocabulario en clase, entre otros.

Los centros de interés seleccionados para las pruebas de Monterrey y Houston son diez para cada idioma, de entre los cuales se incluyen cuatro centros presentes en el estudio pionero en Francia (Gougenheim *et al.*, 1964) y en la investigación escocesa de referencia (Dimitrijević, 1969). En español son los siguientes: 1) *Animales*; 2) *La ciudad*; 3) *Diversión*; 4) *Familia*; 5) *Dios*; 6) *Trabajos y profesiones*; 7) *Medicina*; 8) *Música*; 9) *Espacio*; 10) *Guerra*. En inglés se corresponde a la traducción literal de esos mismos centros: 1) *Animals*; 2) *The city*; 3) *Entertainment*; 4) *Family*; 5) *God*; 6) *Jobs and professions*; 7) *Medicine*; 8) *Music*; 9) *Space*; 10) *War*. La recogida de datos se realiza mediante listas abiertas donde a cada centro se le asigna una duración de dos minutos⁶².

Los resultados muestran que los hablantes monolingües de inglés ocupan el primer rango en cuanto a productividad ya que aportan 6440 unidades léxicas. En el siguiente puesto se sitúan los hablantes bilingües de español con un total de 5672 palabras, seguidos de los hablantes bilingües de inglés con 5572 palabras y, por último, los hablantes monolingües de español con 4696 palabras. Sin embargo, los datos difieren al contabilizar el número de vocablos ya que los hablantes bilingües de español son los que muestran más diversidad léxica con 2539 palabras diferentes, seguidos de los hablantes monolingües de inglés con 2454 vocablos, los hablantes bilingües de inglés con 2384 y los hablantes monolingües de español con 1904 vocablos.

Para llevar a cabo un análisis adecuado, el estudioso americano relaciona los resultados con otras variables, como el estatus sociocultural. Por ejemplo, basándose en el estatus ocupacional de las familias y los distritos donde se sitúan sus casas, deduce que los hablantes monolingües de español tienen un estatus superior al de los hablantes monolingües de inglés y los hablantes bilingües. El autor concluye que las diferencias lingüísticas se deben de forma general al nivel socioeconómico, al sexo y la

⁶² The testing was divided into distinct areas of subject matter and for each, the students gave, by means of a free association type response, words which they related directly or indirectly to one stimulus at a time. The participants were required to write down their responses as pre-testing experimentation uncovered some severe disadvantages in using oral-type recording devices. Each stimulus was allowed two minutes (Victory, 1971: 1).

edad, y más específicamente a su estatus, inteligencia, interés personal, aptitudes y actitudes.

En 1971 y 1979 Galisson reflexiona sobre la efectividad pedagógica de los listados del *Français Fondamental* y realiza aportaciones metodológicas que, a su juicio, complementan los estudios de disponibilidad léxica. Desde el punto de vista de un docente, considera que los listados ordenados alfabéticamente suponen un esfuerzo mayor a la hora de localizar una palabra que quiera enseñar frente a listas de palabras organizadas por temáticas. Además, ese vocabulario no responde a la lengua estándar que se pretende transmitir ni comprende todas las palabras gramaticales necesarias. Por tanto, los centros de interés llevarían a una serie de sustantivos y verbos independientes de utilidad dudosa.

La propuesta de Galisson conlleva la realización de dos pruebas complementarias: una lexemática, donde a partir de un tema determinado se favorezca la producción de sustantivos y verbos y, otra de tipo sintagmático, donde se registren palabras asociadas con los campos dados. Por tanto, la prueba de la disponibilidad léxica, debería llevar anexa otra de tipo asociativo que consiguiese captar los vocablos más disponibles referidos a otro tipo de palabras gramaticales como los adjetivos y adverbios.

Además, incluye el concepto de *temas predilectos* en contraposición con los centros de interés. Esos temas suponen la reflexión y asociación léxica, directa o indirecta, en torno a un campo determinado de la experiencia, mientras que los centros de interés solo muestran las palabras que pertenecen a una categoría. Los temas predilectos en torno a los que organiza el inventario temático y sintagmático son los siguientes: 1) *Maison*; 2) *Appartement*; 3) *Repas*; 4) *Ménage*; 5) *Famille*; 6) *Corps*; 7) *Vêtements*; 8) *Soins*; 9) *Ville*; 10) *Commerce*; 11) *Voyage*; 12) *Poste*; 13) *Identité*; 14) *École*; 15) *Clase*; 16) *Carte*; 17) *Temps*; 18) *Travail*; 19) *Agricultura*; 20) *Gouvernement*; 21) *Plaisirs*; 22) *Religion*; 23) *Relations humaines*. Esa clasificación permite al hablante no solo conocer el significado y el uso de las palabras en contexto, sino que también soluciona problemas relacionados con la presencia de las palabras homónimas y polisémicas.

Posteriormente, Galisson (1991) realiza una clasificación de los tipos de palabra desde el punto de vista de la didáctica de lenguas extranjeras o segundas lenguas. Distingue tres tipos de palabras: coocurrentes, correlacionadas y colaterales. Las primeras hacen referencia a los vocablos relacionados con una palabra concreta pero de categoría sintáctica diferente, así algunas palabras coocurrentes de *agua* podrían ser *beber*, *bañarse* y *húmedo*. Las palabras correlacionadas pertenecen a la misma categoría gramatical y guardan alguna relación (sinonimia, antonimia e hiperonimia) que permite su intercambio en un contexto dado. Por ejemplo, la palabra *guapo* está correlacionada con *feo*, *bello*, *bonito*. Por último, las palabras colaterales serían las unidades léxicas de la misma familia, como *pan* y *panadería*.

La disponibilidad léxica también viaja a África donde se analizan los resultados en una o dos lenguas diferentes con fines pedagógicos y, sobre todo, sociológicos (Fall, 1976; Njock, 1978; Gontier, 1986). Los primeros intentos de aplicar la metodología francesa remontan al año 1966, cuando el Centre de Linguistique Appliquée de Dakar decide llevar a cabo la elaboración de un vocabulario fundamental del wolof. El objetivo es conseguir un corpus de 5000 palabras mediante dos métodos de recogida de datos tradicionales: entrevistas orales y cuestionarios de disponibilidad léxica.

Las entrevistas se realizaron a 1000 senegaleses sobre temas familiares durante 15 minutos. En total se registraron 200 horas de grabaciones y se seleccionaron para su posterior análisis las 200 000 primeras palabras. En cambio, la prueba de disponibilidad léxica se diseñó pero nunca llegó a realizarse. El centro de lingüística había decidido limitar el tiempo de los centros de interés a 10 minutos, seleccionar como muestra de estudio a hablantes 18 años o más y solicitar datos relativos a su origen regional, social, nivel educativo, profesión, sexo y edad.

En 1974 Fall comienza a realizar los primeros sondeos con alumnos de escuelas de primaria de Dakar y alrededores. En 1976, presenta *L'univers Culturel de l'enfant wolophone de Dakar, d'après la disponibilité des concepts dans la langue maternelle et la langue d'enseignement (cas du wolof et du français)*. Tal y como indica el título, el estudio pretende acercarse al universo cultural del niño que habla wolof a través de una prueba de asociaciones léxicas, como es la disponibilidad léxica. Los centros

seleccionados para la prueba son los siguientes: 1) *Les parties du corps*; 2) *La nourriture et les boissons*; 3) *Les vêtements*; 4) *Les jeux et les distractions*; 5) *Les animaux*; 6) *Les métiers*; 7) *La maison et son mobilier*; 8) *Les moyens de transport*; 9) *La ville*; 10) *La cuisine*. La utilidad de este trabajo es incontestable en la fabricación de métodos de alfabetización y de programas de educación popular.

Años más tarde, Gontier (1986) presenta su tesis doctoral titulada *Exécution, traitement, analyse et commentaire d'une enquête de disponibilité lexicale en milieu scolaire sénégalais (classes de CM2, 5^{ème} et 3^{ème})*⁶³. En este caso, se analizan los datos solamente en francés de estudiantes de tres cursos diferentes⁶⁴. Cabe destacar que anteriormente Gontier publica junto con Blondé y Dumont (1979a) el léxico del francés de Senegal que consiste en un inventario donde añade a las palabras del francés neutro una cita o comentario donde se reflejan las particularidades del francés senegalés, a menudo influenciado por el wolof. De forma complementaria Gontier, Blondé y Dumont (1979b) presentan en otra obra las particularidades del francés de Senegal.

En esa misma línea, se encuadra el trabajo de Njock (1978)⁶⁵ que hace referencia al contexto social del niño africano más que a su lengua o léxico. El objetivo pasa por conocer el medio familiar del niño basaa de Camerún por medio de la lengua. La prueba de disponibilidad léxica permite conocer las palabras disponibles que vienen a la mente al individuo cuando habla o escribe. Según el autor, esas palabras parecen pertenecer a una serie de asociación de ideas que solo aparecen cuando una serie de ese tipo entra en juego (Njock, 1978: 47-60). Para argumentar este enfoque, el investigador realiza una descripción pormenorizada del proceso de asociación de palabras.

⁶³ Citado por Rodríguez en *Cahiers franco-canadiens de l'Ouest* (2007) como «Gontier, D. (1986) *Le vocabulaire disponible des enfants sénégalais, thèse* (doctorat de 3^e cycle), Université de Strasbourg II, 226 p». En cambio, en la mayoría de revistas aparece citado como «Gontier, D. (1986): *Exécution, traitement, analyse et commentaire d'une enquête de disponibilité lexicale en milieu scolaire sénégalais (classes de CM2, 5^{ème} et 3^{ème})*, Thèse de troisième cycle, Université de Strasbourg II», por ejemplo en la *Revue du Réseau des Observatoires du Français Contemporain en Afrique Noire. Le Français en Afrique*. Disponible en <http://www.unice.fr/bcl/ofcaf/11/11G.html> [Fecha de última consulta el 6 de abril de 2015].

⁶⁴ Cours Moyen 2^{ème} année, corresponde a 5º de Primaria, 5^{ème} equivale a 1º de la ESO y 3^{ème} a 3º de la ESO.

⁶⁵ En la cubierta aparece el año 1979 pero la publicación tiene lugar el año anterior, 1978. La obra original está disponible en <http://0-files.eric.ed.gov.opac.msmc.edu/fulltext/ED192580.pdf> [Fecha de última consulta el 5 de marzo de 2015].

En effet le lexique se rattache de très près aux réalités socio-culturelles. Pour mieux couvrir ce milieu socio-culturel, nous avons choisi d'interroger l'enfant dans sa langue maternelle et dans la langue d'enseignement: le basaa et le français tels qu'il sont parlés dans deux départements du Cameroun⁶⁶ (Njock, 1978: 45).

No obstante, el autor señala ocho factores que condicionan la disponibilidad léxica y, por tanto, la asociación de ideas en la mente de los informantes: los centros de interés elegidos; la época, es decir, del momento de la realización de las pruebas; la estación del año⁶⁷; el medio socioeconómico; el sexo; el sector educativo (escuela técnica, clases de enseñanza general, establecimiento privado, laico, etc.); el nivel, es decir el nivel de enseñanza y el grado de escolarización del niño y la edad. Por ello, las pruebas se realizan sobre una muestra estratificada y demográficamente representativa de la población entera: 220 alumnos bilingües (basaa-francés) escolarizados en un nivel educativo de primaria.

Los datos se han recopilado en las dos lenguas sobre los siguientes diez centros de interés: 1) *Les parties du corps*; 2) *La nourriture et les boissons*; 3) *Les vêtements (pour hommes et pour femmes)*; 4) *La maison: matériaux de construction et meubles*; 5) *Les travaux des champs*; 6) *La cuisine*; 7) *La salle de classe*; 8) *Les moyens de locomotion*; 9) *La musique et la danse - les genres*; 10) *Les jeux et divertissements*. Cabe destacar que en el centro de interés relativo a los juegos y las distracciones los niños aportan muchas más palabras en basaaa que francés e incluso más que los jóvenes franceses en su lengua.

El estudio registra en total 88 000 ocurrencias que son estudiadas, clasificadas y comparadas. Los análisis de palabras se dividen en cinco grupos: palabras disponibles en lengua materna, palabras disponibles en la lengua de enseñanza, palabras disponibles en el mismo nivel en las dos lenguas, palabras disponibles en las dos lenguas en diferentes grados y palabras de los informantes bilingües. Además, la investigación da cuenta de los matices culturales presentes en los dos idiomas: palabras que designan en ambas lenguas la misma realidad, palabras polisémicas en una cultura pero no en otra, palabras con un mismo significante pero diferente significado y viceversa. Del análisis se desprende que hay más similitudes entre el

⁶⁶ Elige la región donde se habla basaa y el pueblo basaa de Camerún.

⁶⁷ Ejemplifica que durante la estación de recogida, la disponibilidad léxica de los vocablos referidos a frutas y verduras será más elevada que si se realizase en otra estación.

basaa y el francés de Camerún que entre el francés de Francia y el francés de Camerún.

La investigación concluye que el bilingüismo supone un factor de enriquecimiento en los hablantes bilingües ya que describen mejor la realidad que los monolingües y, por tanto, son capaces de comprender mejor el medio sociocultural en el que viven. En cuanto al francés de Francia, considera que es inadecuado para describir la realidad sociocultural africana. Sin embargo, resalta la tendencia de las lenguas africanas a utilizar préstamos de otras lenguas en lugar de potenciar la creación léxica con términos nuevos locales para describir la África moderna y sus realidades.

La primera investigación en incluir la disponibilidad léxica en lengua española, además de en inglés, corresponde a la obra de Victory (1971). A continuación, le siguen investigaciones realizadas en Puerto Rico (López Morales, 1973, 1978, 1979; Román, 1985), en México (Justo Hernández, 1986⁶⁸; Cañizal Arévalo, 1987, 1991; Ruiz Basto, 1987); en España (Azurmendi Ayerbe, 1983; Benítez Pérez, 1992) y en Chile (Echeverría, Herrera, Moreno y Pradenas, 1987; Mena Osorio, 1986), entre otros.

No obstante, la metodología seguida por los estudios hispanos difería de un investigador a otro lo que dificultaba la comparación intermuestral (centros de interés, tiempo de la prueba, lematización de las palabras, etc.). Por ello, durante el XI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL) que tuvo lugar en Las Palmas de Gran Canaria, expertos en el ámbito de la disponibilidad léxica decidieron consensuar algunas pautas generales procesuales y analíticas. No será hasta 1999 cuando tenga lugar en la Universidad del País Vasco la reunión científica que establecerá los objetivos y la metodología del Proyecto Panhispánico de Disponibilidad Léxica. El objetivo del Panhispánico, coordinado por López Morales, es conocer el léxico disponible de los diferentes países hispanohablantes a través de diccionarios de disponibilidad creados bajo los

⁶⁸ A diferencia de la tendencia de otras investigaciones, el investigador mexicano trabaja con listas cerradas de diez palabras por centro.

mismos principios. De forma general, las bases sobre las que se asientan son las siguientes:

- La recogida de datos se realiza mediante una encuesta de léxico disponible que utiliza los dieciséis centros de interés tradicionales: 1) *El cuerpo humano*; 2) *La ropa*; 3) *Partes de la casa (sin los muebles)*; 4) *Los muebles de la casa*; 5) *Alimentos*; 6) *Objetos colocados en la mesa para la comida*; 7) *La cocina y sus utensilios*; 8) *La escuela: muebles y materiales*; 9) *Calefacción, iluminación y medios de airear un recinto*; 10) *La ciudad*; 11) *El campo*; 12) *Medios de transporte*; 13) *Trabajos del campo y del jardín*; 14) *Los animales*; 15) *Juegos y distracciones*; 16) *Profesiones y oficios*.
- El tiempo de la prueba está limitado a dos minutos por centros de interés.
- Las listas son abiertas donde el sujeto puede escribir todas las palabras que le vengan a la mente.
- La población objeto de estudio está formada por estudiantes de nivel preuniversitario (concretamente en el curso anterior que da acceso a la universidad). De esta forma, se pretende homogeneizar la edad, el nivel de instrucción de los hablantes e incluso el léxico ya que se considera que esos estudiantes no han adquirido los términos técnicos con estudios superiores.
- La edición de los datos atiende a un protocolo previamente acordado donde se incluyen directrices para la transcripción de los materiales, lematización de sus palabras y eliminación de ciertos vocablos⁶⁹.
- La fórmula matemática recomendada para calcular la disponibilidad léxica es la elaborada por López Chávez y Strassburger Frías (1991).
- Las variables utilizadas corresponden al sexo, nivel sociocultural, zona geográfica (urbana o rural) y tipo de centro (público o privado).
- Los listados se publican en forma de diccionarios, usualmente ordenados por orden alfabético y por orden decreciente de disponibilidad, pero también puede estar organizado según variables determinadas, como el sexo o el estatus social, entre otras.

Con objeto de coordinar el trabajo de los diversos estudios disponibles, se celebraron otros tres encuentros: el II y III Encuentro Internacional de

⁶⁹ Una de las obras de referencia en este sentido es Samper Padilla (1998) y Samper *et al.* (2003: 95-101).

Disponibilidad Léxica se realizaron en el año 2003 y 2005 en San Millán de la Cogolla (La Rioja) y el IV Encuentro de Disponibilidad Léxica tuvo lugar en Salamanca⁷⁰ en 2011. A lo largo de estos años, se han llevado a cabo numerosos trabajos aplicando esta metodología no solo en estudios de disponibilidad léxica en español a hablantes de lengua materna sino también a estudiantes de español como lengua extranjera y hablantes y estudiantes de otras lenguas⁷¹.

2.4. Investigaciones relacionadas con el léxico disponible

No cabe duda que aunque el francés diera vida al campo de la disponibilidad léxica, el español es la lengua en la que se han desarrollado la mayoría de los estudios sobre todo gracias al empuje recibido por el Proyecto Panhispánico de Disponibilidad Léxica. En los apartados anteriores se ha realizado un recorrido por las primeras investigaciones y se han abordado los principales cambios metodológicos desde sus orígenes hasta la actualidad. El objetivo de este apartado es encuadrar el actual estudio en las investigaciones de español como lengua materna y en inglés o francés como lengua extranjera.

En primer lugar, se mencionan las principales investigaciones en lengua española partiendo de Puerto Rico, país donde López Morales comienza realiza su primera investigación sobre disponibilidad léxica, pasando por otros países hispanoamericanos hasta llegar a España, donde los estudios se dividen por comunidades autónomas. En cuanto al español como lengua extranjera o segunda lengua, cabe mencionar el florecimiento de las investigaciones en este ámbito en los últimos años. El precursor, Carcedo González (2000c), analiza el léxico disponible de español como lengua extranjera de estudiantes fineses. Por las características del presente estudio, se han recopilado las principales investigaciones en inglés como lengua extranjera.

⁷⁰ Agradezco tanto a los organizadores del mencionado encuentro como a los directores del presente trabajo que permitieran mi asistencia.

⁷¹ El equipo de DispoLex recoge las referencias bibliográficas de más de 350 estudios publicados sobre la disponibilidad léxica y el directorio con los investigadores principales en este ámbito. [Consultado en www.dispoLex.com el 19 de agosto de 2014].

En cambio, en francés como lengua extranjera no se ha dedicado ningún apartado específico debido al escaso número de investigaciones. Aparte de los estudios de Mackey en Canadá (1971a, 1971b), se han localizado investigaciones en la provincia canadiense de Manitoba, zona donde el inglés es el idioma oficial *de facto* y el francés es el idioma oficial *de iure*. Las principales indagaciones han sido llevadas a cabo por la investigadora Rodríguez, especializada en los estudios de variación geolingüística en Canadá (Rodríguez, 1993, 2004, 2006, 2007, 2009, 2010). Asimismo, Dupré (2007) realiza un estudio comparativo sobre estudiantes bilingües en la región de Manitoba (Saint-Lazare) y Ontario (Ottawa) y analiza más en detalle la disponibilidad léxica de los alumnos en Ottawa en su tesis de doctorado (2002) bajo la dirección de Bougy y Rodríguez. En cuanto a las antiguas colonias francesas, destacan los estudios de Fall (1976) y Gontier (1986) en Senegal y Njock (1978) en Camerún.

2.4.1. En español como lengua materna

2.4.1.1. En Hispanoamérica

A pesar de que la primera investigación en lengua española la llevase a cabo Victory (1971), por la envergadura de los trabajos de López Morales se le considera el precursor de la disponibilidad léxica en español. Asimismo, Victory (1971) trabaja con una muestra variada (alumnos monolingües en México, alumnos bilingües y monolingües de Estados Unidos) y analiza la disponibilidad léxica en español y en inglés mientras que López Morales (1973) realiza un estudio sociolingüístico sobre una muestra de 63 escolares de dos sociolectos de San Juan (Puerto Rico). Los centros de interés seleccionados son los siguientes: 1) *Alimentos*; 2) *Animales*; 3) *La casa*; 4) *La cocina*; 5) *El cuerpo humano*; 6) *Materiales de construcción*; 7) *Muebles*; 8) *La ropa*; 10) *Transportes*.

López Morales (1978) utiliza ese léxico de los escolares de la región metropolitana de San Juan de Puerto Rico recogido en el estudio de 1973 para compararlo con el léxico frecuente de preescolares puertorriqueños llevado a cabo por Rodríguez Bou en 1952. Tras cotejar las dos obras (López Morales, 1973 y

Rodríguez Bou, 1952), López Morales (1978) manifiesta la gran diferencia encontrada en ambos listados no solo por la diferencia de edad de los informantes estudios y su consiguiente tiempo de aprendizaje sino también por la metodología empleada para crear el léxico frecuente y el léxico disponible.

En esa misma línea, López Morales continua con el estudio de la disponibilidad desde dos vertientes: la sociolingüística y la lingüística aplicada a la enseñanza de lenguas. Por un lado, en 1979 se sirve del índice de disponibilidad para analizar y comparar cualitativa y cuantitativamente el léxico disponible de tres estratos sociales (bajo, obrero y medio) de la muestra del estudio San Juan de Puerto Rico de 1973. Por otro lado, en 1986 destaca la importancia del léxico disponible en la selección del vocabulario. Su propuesta pasa por incorporar a los listados de léxico frecuente los resultados del léxico disponible.

Ya en 1991⁷², López Morales destaca la importancia que había cobrado la disponibilidad léxica en los años precedentes. Con objeto de seguir impulsando los estudios en este ámbito, López Morales y García Marcos (1995) coordinan el proyecto de investigación de disponibilidad léxica en Andalucía del que surgieron trabajos en Almería (Mateo García, 1996, 1998) y en Cádiz (González Martínez, 2002), entre otros. Un año más tarde, López Morales (1996) hace una revisión de los primeros trabajos en este ámbito, de la forma de calcular la disponibilidad léxica y de las principales aplicaciones de la disponibilidad léxica en otras disciplinas tales como, la dialectología, la sociolingüística, la psicolingüística, la etnolingüística y la lingüística aplicada a la enseñanza de la lengua.

En 1999, López Morales estudia el léxico disponible de Puerto Rico a partir de una muestra de 558 estudiantes universitarios de primer año. La prueba incluye los 16 centros del *Français Fondamental* aunque con ligeras modificaciones para adaptarlos al clima del país y a la muestra de estudio: *chauffage* se sustituye por *aire acondicionado* y *école* por *universidad*. Con objeto de realizar un estudio sociolingüístico, la prueba

⁷² «Lo relativo a la disponibilidad léxica ha experimentado un auge sin precedentes desde la presentación de López Morales en 1978. En aquel momento, la disponibilidad sólo se trabajaba a base de frecuencias, siguiendo el patrón de análisis establecido por los iniciadores franceses (Michéa 1957, Gougenheim 1967); así había sido elaborado el Proyecto de disponibilidad léxica en escolares de San Juan (López Morales 1973) del que sólo un par llegaron a la letra impresa (López Morales 1978, 1979)» (López Morales: 1991: 12).

incorpora variables sociológicas referidas al sexo, nivel sociocultural, zona geográfica y tipo de escuela donde los informantes estudiaron antes de ir a la universidad.

Tomando como base los datos recopilados en ese estudio, en 1999b el investigador aborda más específicamente los anglicismos presentes en esos listados. Asimismo, otros autores comparan los listados puertorriqueños con distintas variedades del español: López Chávez (1992, 1995) realiza el cotejo con el inventario de Madrid, Las Palmas, República Dominicana y Puerto Rico; Alba (1998) lleva a cabo una comparación dialectal que incluye los listados de República Dominicana, Madrid, México, Concepción (Chile) y Puerto Rico; González Martínez (2000-2001) compara el léxico puertorriqueño con el léxico gaditano; Carcedo González (2003) añade a los listados de Puerto Rico y Cádiz el de Asturias, etc.

Tras la merecida introducción de los trabajos de López Morales y siguiendo la demarcación geográfica, se comentan algunos de los principales trabajos llevados a cabo en América, por países y siguiendo el orden de publicación del primer trabajo en el ámbito. En España, debido a la vasta literatura en este campo, los estudios se presentan según comunidades autónomas.

En Puerto Rico, Belén Román (1985) realiza su tesina sobre una muestra de 293 estudiantes de primer, tercer y quinto grado del Distrito Escolar de Dorado utilizando los 10 centros de interés presentes en López Morales (1973). Entre las conclusiones del trabajo, cabe destacar la relación que guardan los resultados de disponibilidad léxica con las variables sociológicas del estudio: grado escolar, sexo, zona de procedencia y nivel socioeconómico.

En 1987, Butrón hace una revisión de la fórmula matemática de la disponibilidad en un estudio en el que analiza el léxico de alumnos de escuela elemental. Posteriormente, Butrón (1989) analiza algunos aspectos sociolingüísticos que influyen en el léxico disponible presentado en 1987. Asimismo, en 1991 incorpora nuevos índices representativos de la disponibilidad léxica: disponibilidad de centros para un grupo, disponibilidad de léxica de un informante y disponibilidad léxica de un grupo.

En 2013, Fuster Pérez examina el léxico de una muestra de 60 estudiantes de primer año de la Universidad de Puerto Rico estratificada según sexo (30 hombres y

30 mujeres) y según el tipo de centro educativo (30 provienen de la escuela pública y 30 de la escuela privada). El estudio recoge tan solo cuatro centros de interés, de los cuales realiza un análisis cuantitativo: 1) *Comidas y bebidas*; 2) *La ciudad*; 3) *Juegos y distracciones*; 4) *Profesiones y oficios*. El análisis cualitativo lo realiza solamente sobre el centro de interés *Juegos y distracciones* el cual compara con el léxico recogido por López Morales 23 años antes, es decir, en 1990⁷³. Cabe mencionar que para la comparación solo se tuvieron en cuenta las 100 unidades léxicas más disponibles ya que el límite de tiempo en López Morales era de cinco minutos y el de Fuster Pérez (2013) dos minutos.

En Chile, el estudio pionero en disponibilidad léxica corresponde a la tesis de maestría realizada por Mena Osorio (1986) a un grupo de 300 escolares de 4.º, 6.º y 8.º grado de educación básica bajo la dirección de Echeverría, considerado el impulsor de la disponibilidad léxica en ese país. Se trata de la primera investigación que establece como límite temporal dos minutos por cada centro de interés y que se ha convertido en la pauta seguida por la mayoría de investigadores.

Echeverría (1987) dirige el equipo de investigación compuesto por Herrera, Moreno y Pradenas que se encarga de realizar la disponibilidad léxica de 200 estudiantes de primer curso y 200 de tercer curso de Educación Media de Concepción. El estudio incluye el cálculo del índice de cohesión, tan extendido en numerosas investigaciones, y divide los centros de interés según ese grado de cohesión semántica, distinguiendo entre centros de alto grado: 1) *Cuerpo humano*, 2) *Medios de transporte*, 3) *La ropa*, 4) *Muebles*, 5) *Herramientas y materiales de construcción*, 6) *Alimentos*, 7) *Animales*; y centros de bajo grado 8) *Naturaleza*, 9) *La casa* y 10) *La cocina*. En 1991, Echeverría realiza un estudio comparativo del léxico de estudiantes de nivel básico y nivel medio recogidos en los estudios anteriormente mencionados (Mena Osorio, 1986 y Echeverría *et al.*, 1987). En el análisis destaca que además del nivel de escolaridad, la variable que más influye en el léxico disponible es la procedencia sociocultural.

Ese mismo año, Vargas Sandoval (1991) pretende conocer el léxico hablado por la clase media de los adultos de la Región de Valparaíso. La muestra de estudio

⁷³ Los datos del estudio de López Morales (1999) fueron recopilados en 1990.

está formada por 500 informantes entre 20 y 50 años de las doce ciudades más pobladas de esa región (San Antonio, La Ligua, Quillota, La Calera, Limache, Llay-Llay, Quilpué, Villa Alemana, San Felipe, Los Andes, Valparaíso y Villa del Mar). Los centros de interés analizados son nuevos: 1) *Arte*; 2) *Derecho*; 3) *Economía*; 4) *Política*; 5) *Sentimiento*; 6) *Teología*.

En 1997 comienza un proyecto de investigación de tres años de duración bajo la dirección de Echeverría destinado a la creación de tres programas informáticos para la enseñanza del español. Concretamente el programa denominado *Vocabulario disponible* trata de buscar una aplicación práctica al léxico disponible dentro de la lingüística aplicada a la enseñanza de una lengua. El léxico incorporado forma parte del estudio de Valencia Espinoza y Echeverría (1999) que analiza el léxico disponible de los estudiantes chilenos de todas las regiones del país. El número de informantes asciende a un total de 2052 escolares de 4.º año de educación media que aportan información sobre el sexo, nivel sociocultural, régimen educativo, región y procedencia poblacional. Trabajan con 18 centros de interés, en su mayoría diferentes a los tradicionales, lo que dificulta la comparación con otros estudios del Proyecto Panhispánico⁷⁴.

A partir de los listados del léxico disponible de Chile, Arancibia Manhey (2006) analiza las unidades léxicas procedentes de lenguas indígenas, concretamente del mapudungun, el quechua, el náhuatl, el caribe, el arahuaco, el kunza, el kuna, el taíno y el guaraní. La investigadora extrae 301 vocablos provenientes de alguna de las 10 familias aborígenes de América y los presenta ordenados alfabéticamente según el centro de interés y con indicación de la lengua de origen y la frecuencia absoluta.

Las lenguas más productivas son el quechua y el mapudungun y las menos el aymara, el kuna y el kunza. La influencia de diferentes culturas queda patente en las unidades léxicas aportadas por los informantes chilenos ya que, por ejemplo, en el centro de interés denominado *Mundo espiritual* aparecen vocablos provenientes de la

⁷⁴ 1) *Partes del cuerpo*; 2) *Ropa y calzado*; 3) *La casa*; 4) *Muebles*; 5) *Alimentos*; 6) *La cocina*; 7) *Medios de transporte*; 8) *Trabajo del campo y del jardín*; 9) *Plantas y animales*; 10) *Juegos y entretenimiento*; 11) *Profesiones y oficios*; 12) *La política*; 13) *La actividad económica*; 14) *Las artes*; 15) *Ciencia y tecnología*; 16) *Mundo espiritual*; 17) *Procesos mentales*; 18) *Problemas del ambiente*.

mitología de la Isla de Chiloé (*caicai-viñú, calenche, pincoya*) y de la zona andina (*Viracocha* y *Mamacocha*).

Según Arancibia Manhey (*cf.* 2006: 71) en los primeros tiempos de la Conquista y la Colonia la lengua española introdujo los vocablos de los productos autóctonos. Por esta razón, la mayoría de las entradas léxicas procedentes de lenguas indígenas que forman parte del acervo lingüístico español se localizan en el centro de interés *Plantas y animales*, seguido de *Alimentos* y *Trabajos del campo y del jardín*. Por ejemplo, en el centro de *Alimentos* aparecen vocablos como *maíz, cacao, mate, chocolate, papa, tomate*; en el centro de *Plantas y animales* se registran las palabras *coca, yerba mate, chirimoya, puma*; en la *Ropa y calzado* se encuentra *alpaca*, etc.

Echeverría, Urzúa e Figueroa (2005) presentan una nueva aplicación informática para calcular el índice de disponibilidad denominada «Dispogen II». Un año más tarde, Urzúa, Sáez y Echeverría (2006) estudian la disponibilidad léxica de estudiantes y profesores de la carrera de Ingeniería Civil Matemática en la Universidad de Concepción. La muestra está compuesta por 85 informantes, 71 estudiantes (divididos en tres niveles según el curso en el que estaban matriculados) y 14 profesores. Los seis centros de interés se adaptan a la muestra de estudio: 1) *Cálculo*; 2) *Álgebra*; 3) *Estadística*; 4) *Física*; 5) *Educación*; 6) *Geometría*. El objetivo es conocer cómo evoluciona la adquisición de ese vocabulario específico en función de los años de estudio en ese ámbito y reflexionar sobre los errores y lagunas léxicas de los estudiantes para intentar tomar medidas al respecto.

Ferreira (2006) presenta su tesis de magíster sobre la disponibilidad léxica en inglés como lengua materna y como lengua extranjera sobre la que más tarde, en el año 2010, publica un artículo junto con Echeverría. En su trabajo destaca la aplicación de la psicolingüística a los estudios de disponibilidad léxica a través del análisis de las relaciones de palabras con el programa DispoGrafo, aplicación presentada años más tarde en un artículo por Echeverría, Vargas, Urzúa y Ferreira (2008)⁷⁵. Se trata de la primera herramienta informática que contabiliza el número de

⁷⁵ El software fue creado por Echeverría y Vargas Opitz en 2006.

relaciones entre dos palabras, los *clusters* y que, además, es capaz de representar de forma gráfica esa organización de palabras en cada centro de interés⁷⁶.

Fasce, Echeverría, Matus, Ortiz, Palacios y Soto (2009) publican un artículo interdisciplinar en la *Revista médica de Chile*. El fin del estudio es conocer los atributos o cualidades que estudiantes de medicina y médicos en ejercicio identifican como recomendables para el ejercicio de la profesión a través de la disponibilidad léxica.

Las investigaciones en México comienzan en 1986 con la tesina de licenciatura de Justo Hernández, *Disponibilidad léxica en colores*. La muestra está formada por 56 informantes adultos (26 mujeres y 30 hombres) y las variables socioculturales incluidas en el estudio son el sexo, el tipo de enseñanza y la ubicación escolar. El principal objetivo consiste en conocer los nombres de los colores más disponibles relacionados con diez centros de interés mediante el sistema de listas cerradas en los que los informantes solo tenían que escribir diez nombres de colores en cada centro⁷⁷.

Cañizal Arévalo (1987) analiza la influencia de ciertos factores sociológicos como el sexo, el tipo de escuela (pública y privada) y las horas de exposición a la televisión en seis centros de interés en una muestra de 68 escolares de primaria. Al igual que Justo Hernández (1986) la autora pretende recoger adjetivos disponibles de forma incluye el centro de interés *Defectos físicos y morales* para obtener mayores resultados. Los otros cinco centros de interés forman parte del listado tradicional: 1) *Partes del cuerpo*; 2) *Objetos que hay en el interior de la casa*; 3) *Alimentos y bebidas*; 4) *Medios de transporte*; 5) *Profesiones y oficios*. En 1991, estudia la organización de las palabras de un centro de interés destacando la existencia de un núcleo en torno al cual se organizan el resto de palabras en forma de redes o constelaciones.

Las contribuciones mexicanas más importantes vienen de la mano de López Chávez y Strassburger Frías (1987, 1991) gracias a la revisión de la fórmula matemática que permite calcular el índice de disponibilidad léxica⁷⁸ que es empleada

⁷⁶ En el apartado 6.5 se explica el funcionamiento del programa con detalle y se presentan los análisis realizados con el programa en este estudio. Agradecemos la colaboración de Roberto Ferreira quien nos facilitó el DispoGrafo y el Dispogen.

⁷⁷ 1) *La ropa*; 2) *Zapatos*; 3) *Cosméticos*; 4) *Automóviles*; 5) *Anuncios comerciales*; 6) *Mochilas, bolsas y maletas*; 7) *Muebles*; 8) *Tapices y alfombras*; 9) *Pintura de casa*; 10) *Vestido*.

⁷⁸ *Vid.* apartado 2.5 para una revisión de la fórmula matemática.

casi de forma generalizada en los estudios posteriores y en las principales aplicaciones informáticas (Lexidisp, Dispolex, Dispogen y Ledical). En 1993, López Chávez analiza el léxico disponible de estudiantes mexicanos desde preescolar hasta sexto grado de la escuela primaria a través de los 16 centros tradicionales. En total distingue siete edades diferentes con objeto analizar la organización de las palabras en la mente y la evolución en cada tramo. Un año más tarde, en 1994, realiza un estudio en la ciudad de México sobre 30 estudiantes de 5.º de Bachillerato y selecciona centros de interés que favorecen la aparición de verbos.

En 1990, Hernández Solís emplea por primera vez en México la fórmula de López Chávez y Strassburger Frías (1987) en un estudio de disponibilidad léxica sobre una muestra reducida de alumnos de 6.º de primaria y 1.º de secundaria provenientes de escuelas privadas. En el año 2000, la investigadora recoge el léxico disponible de 540 estudiantes de primaria de la ciudad de Zacatecas provenientes de diferentes escuelas (particular, estatal y federal). La muestra estaba previamente estratificada de forma se seleccionaron 30 informantes (15 hombres y 15 mujeres) por cada grado escolar de primaria (de 1.º a 6.º) y por cada tipo de escuela. Por tanto, entre las variables sociológicas se incluye el sexo, la edad, el grado, el tipo de escuela y el nivel socioeconómico. Al igual que en la investigación anterior, selecciona los dieciséis centros de interés tradicionales.

Hernández Solís (2006) presenta la tesis doctoral *Un modelo para la planificación de la enseñanza del vocabulario con fundamento en el léxico disponible de universitarios zacatecanos* bajo la dirección de López Chávez. Como su nombre indica, el estudio tiene como objeto proporcionar orientaciones metodológicas para la planificación de la enseñanza del léxico. Para ello, realiza un análisis sobre el léxico de los universitarios del estado de Zacatecas ya que considera que aportan las unidades léxicas que se deben enseñar en los niveles preparatorios. Asimismo, compara ese léxico con el obtenido en las pruebas de los escolares y el proveniente de la base estadística del *Diccionario del Español de México* (DEM). Por último, realiza una propuesta didáctica en la que expone los componentes de una lección centrada en el léxico aunando así la teoría de la disponibilidad léxica con su aplicación práctica.

En la República Dominicana, Alba Ovalle (1993) realiza un estudio sociolingüístico en una muestra de 65 informantes universitarios y de educación secundaria sobre seis centros de interés: 1) *Alimentos*; 2) *Juegos y diversiones*; 3) *La escuela*; 4) *Medios de transporte*; 5) *Profesiones y oficios*; 6) *El vestido*. El objetivo es conocer la influencia de ciertas variables, referidas al nivel socioeconómico y nivel de escolaridad, sobre el léxico disponible. Algunos datos de su trabajo los presenta en el X Congreso Internacional de la ALFAL celebrado en México en el año 1993 donde compara el léxico disponible de los alumnos dominicanos universitarios y de la etapa de educación secundaria desde una perspectiva sociolingüística.

En 1995, Alba Ovalle analiza el léxico disponible de una muestra de 347 estudiantes dominicanos de primer año universitario que realizan la prueba de léxico disponible sobre los 16 centros tradicionales. La recogida de datos se configura mediante el sistema de listas abiertas con una limitación temporal de dos minutos. El índice de disponibilidad se calcula a partir de la fórmula de Chávez y Strassburger e incorpora el índice de cohesión de Echeverría *et al.* (1987).

El listado resultante ha servido de punto de partida para análisis posteriores donde se ha comparado con otras variedades lingüísticas: López Chávez (1992) compara los léxicos de Madrid, Las Palmas de Gran Canaria, República Dominicana y Puerto Rico; Carcedo González (2000b) contrasta el léxico dominicano con los listados de Cádiz, Chile, Madrid, Puerto Rico y Zamora o incluso con el léxico de alumnos fineses (2000c); Galloso Camacho (2002a) parangona el léxico de la República Dominicana con otras sintopías provenientes de Ávila, Salamanca y Zamora; Samper *et al.* (2003) ordenan los promedios por rangos de los estudios de Puerto Rico, República Dominicana, Madrid, Gran Canaria, Almería, Cádiz, Córdoba, Bilbao, Ávila, Salamanca, Zamora y Asturias.

Además, Alba Ovalle (1999b y 200a) analiza los anglicismos presentes en los listados del léxico dominicano haciendo y presenta su densidad en el léxico dominicano. En este sentido, Bartol Hernández (1998) realiza un estudio comparativo de los anglicismos en República Dominicana y Zamora con los datos de Alba Ovalle (1995b) y Galloso Camacho (1998). Casanova Ávalos (2008) también compara los anglicismos de la República Dominicana con otros estudios de Puerto

Rico, Cádiz, Zamora y Castellón. En 1998 y 2000b, Alba Ovalle compara el listado de la República Dominicana con el de Madrid, México y Concepción donde coteja las cincuenta palabras más disponibles de tres centros de interés: 1) *El cuerpo humano*; 2) *Medios de transporte*; 3) *Alimentos*.

En el año 2011, Alba Ovalle participa en el XVI Congreso Internacional de la ALFAL⁷⁹ donde presenta los datos de un estudio diacrónico que compara los resultados de los listados de léxico disponible recopilados en 1989 y 2008. En ambas muestras se constata que a medida que aumenta el nivel sociocultural, los informantes aportan un mayor número de palabras y viceversa. Asimismo, comenta que esas diferencias se han acentuado en la prueba de 2008.

En Costa Rica, Murillo Rojas realiza su tesis de licenciatura en 1991 sobre el léxico disponible de preescolares y analiza tres variables sociolingüísticas: sexo, ocupación del padre y lugar de procedencia del niño. En 1993 presenta los datos de esa misma población de estudio referidos a cinco centros de interés y un año más tarde, en 1994, profundiza en el análisis del centro de interés *Comidas y bebidas*. En 1998, Murillo Rojas analiza la productividad léxica de una muestra más amplia: niños de preescolar y de primer ciclo de la educación básica. Además, publica de forma conjunta con Sánchez Corrales un artículo sobre los campos semánticos en los preescolares en 1993 y el listado del léxico disponible en 1999⁸⁰.

En el año 2006 Sánchez Corrales y Murillo Rojas presentan la *Disponibilidad léxica de los niños preescolares costarricenses*. La muestra estratificada selecciona 108 niños de kínder provenientes 19 subregiones educativas del país, de colegios públicos y privados y de ambos sexos. Los centros de interés seleccionados eran tan solo 10⁸¹, por lo que recientemente, ambos investigadores están recopilando datos de su país siguiendo las recomendaciones del Proyecto Panhispánico.

En Colombia, se ha realizado una encuesta de disponibilidad léxica en Popayán, llevada a cabo por Bonilla Rengifo (2004) en su tesina realizada en la

⁷⁹ El título de la comunicación es «Evolución del léxico disponible de los dominicanos: observación del cambio en tiempo real» pero no se recoge en las actas del congreso. Los resultados de esa ponencia han sido publicados con posterioridad, en 2013.

⁸⁰ Sánchez Corrales y Murillo Rojas en 2005 amplían el análisis al léxico básico, oral y escrito, y publican el artículo titulado «Elementos diintegrativos del léxico básico escolar costarricense».

⁸¹ 1) *Partes del cuerpo*; 2) *Vestido*; 3) *Juegos y juguetes*; 4) *Alimentos*; 5) *Familia*; 6) *Animales*; 7) *Profesiones y oficios*; 8) *Comunidad*; 9) *Medios de comunicación*; 10) *Medios de transporte*.

Universidad de Salamanca. En la capital, Bogotá, el grupo de investigación formado por Henríquez Guarián, William Santiago, Ruiz, Mateus Ferro está trabajando en la recopilación de material. En la zona Caribe, el grupo de investigación GIESCA, dirigido por Rodríguez, pretende publicar el léxico disponible del Caribe. En esa zona algunos investigadores como Wandy Valdez han dirigido estudios exploratorios como el trabajo de pregrado de Rodríguez Cadena (2007) sobre la disponibilidad léxica de los estudiantes de educación media en Barranquilla.

En Argentina, Wingeyer (2007) realiza su tesis doctoral sobre el léxico disponible de la ciudad de Resistencia sobre una muestra de 250 estudiantes de primer curso de escuela media y 242 estudiantes de último curso. El estudio incluye los dieciséis centros de interés clásicos y añade *Los colores* para analizar el progreso de los estudiantes durante la etapa en la escuela media. Actualmente, Wingeyer dirige el trabajo de Angelina Oliveira, sobre el léxico disponible de Corrientes, que utiliza los mismos centros de interés que en su estudio además de centros particulares, como: *¿Qué hacemos para divertirnos los fines de semana a la noche?*, *Mitos populares*, *¿Cómo insultamos?* y *Formas de expresar alegría ante cosas que nos comunican los demás*. En este caso, los 500 informantes son estudiantes de último curso de la escuela secundaria.

Wingeyer y Angelina Oliveira (2008) analizan las unidades léxicas presentes en el centro de interés *El campo* de los estudios de Corrientes y Resistencia. En esa misma línea, Wingeyer presenta en 2009 dos comunicaciones en el I Congreso Internacional Léxico e Interculturalidad⁸². La primera junto con Gusberti (Wingeyer y Gusberti, 2012) se centra en el léxico disponible relacionado con los mitos populares de Resistencia y Corrientes y, la segunda con Angelina Oliveira (Wingeyer y Angelina Oliveira, 2012) analiza las unidades léxicas que se usan para insultar en Corrientes. Colombo, Trevisán y Wingeyer (2011) también se sirven de esos listados del léxico disponible y analizan el centro de interés *Partes del cuerpo* de Resistencia y Corrientes y en 2013 publican un artículo sobre las representaciones de la sexualidad en el léxico disponible de la región de NEA.

⁸² El congreso tuvo lugar del 11 al 14 de agosto de 2009 en la Universidad Nacional de Tucumán (Argentina) y las actas se publicaron no se publicaron hasta el año 2012.

Saine Camargo (2008) realiza la tesis doctoral sobre el léxico disponible de Buenos Aires. La muestra de estudio está formada por 225 estudiantes de escuelas públicas y 140 de escuelas privadas de secundaria entre 17 y 19 años provenientes de las quince localidades de La Matanza, uno de los 24 partidos del conurbano bonaerense. Además de los centros tradicionales, incorpora cuatro diferentes: *Defectos y virtudes*, *La inseguridad*, *La lectura* y *Tipos de baile*. El estudio del léxico disponible va acompañado de un análisis de corte sociolingüístico según las variables tipo de centro, sexo y nivel sociocultural. Como se puede apreciar, los estudios en ese país son recientes pero muy prósperos. Asimismo, existen otras investigaciones que se están llevando a cabo en este momento como el estudio del léxico en el habla de santarroseño, por medio del grupo de investigación dirigido por Barroso y Forte, o la investigación en Córdoba de Dalurzo.

En Guatemala, Montenegro (2010) realiza un informe sobre el léxico disponible del departamento de Guatemala. La muestra está formada por 200 estudiantes de tercero y 200 estudiantes de sexto de primaria de escuelas públicas donde se demuestra que los estudiantes de sexto de primaria un mayor caudal léxico. La prueba se realiza sobre seis centros de interés: 1) *Partes del cuerpo*; 2) *Animales*; 3) *Alimentos*; 4) *Medios de transporte*; 5) *Profesiones y oficios* y 6) *Accidentes geográficos*. En el estudio se destacan palabras coloquiales y guatemaltequismos y se hace hincapié en la puesta en práctica de los resultados en la planificación del área curricular Comunicación y lenguaje L1 o L2 en tercero y sexto de primaria.

En otros países hispanoamericanos se están llevando a cabo investigaciones que pretenden contribuir al Proyecto Panhispánico. En Uruguay, Gayne ha recogido las muestras de disponibilidad léxica; da Silva Rodríguez completa las encuestas en la zona fronteriza con Brasil; Ramírez en Paraguay; Vásquez en Panamá; Maldonado, Canales y Herranz en Honduras; Gregory en Cuba; entre otros.

2.4.1.2. En España

España es, sin duda, el país donde se han realizado el mayor número de investigaciones relacionadas con el léxico disponible. A continuación, se comentan las principales investigaciones en este territorio haciendo alusión al estado del

Proyecto Panhispánico. El primer estudio, es llevado a cabo por Azurmendi Ayerbe en 1983 con el objetivo de hacer una indagación sociolingüística en el País Vasco sobre el bilingüismo en euskera y español. La recogida de datos la realiza con modificaciones respecto a las pautas de Mackey (1971a, 1971b) ya que utiliza el sistema de listas cerradas de 20 palabras por cada centro de interés en un tiempo límite de 10 minutos (el investigador canadiense concedió 15 minutos).

Aunque los informantes aportaron datos sobre quince centros de interés⁸³, la autora analiza cinco: *Partes del cuerpo*; *Prendas de vestir de calzar*; *Juegos, diversiones y entretenimientos*; *Agricultura y vegetales, ganadería y animales, pesca y peces*. La muestra está compuesta por 1224 escolares de 5.º y 8.º de EGB de la comarca de San Sebastián. El cálculo del índice de disponibilidad lo realiza a partir del cómputo de la frecuencia, al igual que el trabajo pionero de Gougenheim *et al.* (1964). Las principales variables sociolingüísticas analizadas son el sexo, la edad, el lugar de residencia del encuestado, el tipo de centro educativo, la lengua materna, la lengua empleada con amigos y la lengua en la que se realiza la encuesta.

Benítez Pérez (1992) podría considerarse el primer investigador que lleva a cabo un estudio de disponibilidad léxica siguiendo la metodología actual del Panhispánico: utiliza el sistema de listas abiertas con un límite de dos minutos, incorpora los dieciséis centros tradicionales y calcula el índice de disponibilidad según la fórmula de López Chávez y Strassburger Frías (1987, 1991). La muestra selecciona a 257 estudiantes madrileños de COU y analiza tres variables sociales: sexo, nivel sociocultural y tipo de educación de los informantes (pública o privada).

A partir de los resultados obtenidos, Benítez Pérez realiza diversas publicaciones donde pone la atención en aspectos diferentes. Desde un enfoque cuantitativo, analiza el léxico disponible según la variable sexo y concluye que los hombres aportan un mayor número de palabras en diez centros de interés (Benítez Pérez, 1994a). Ese mismo año, compara los resultados de los listados de

⁸³ 1) *Partes del cuerpo*; 2) *Prendas de vestir y de calzar*; 3) *Partes u objetos que hay en la vivienda (casa o piso)*; 4) *Alimentos y bebidas*; 5) *Relaciones familiares (madre, tío) y no familiares (alumno, amigo)*; 6) *Ciudad o pueblo: sus partes (lo que se ve y se hace en ellos)*; 7) *Los transportes: medios y vehículos que se utilizan*; 8) *Enseñanza: clases de centros y sus partes, objetos y materiales utilizados, asignaturas, etc.*; 9) *Trabajo: clases de trabajos, de oficios, de profesiones*; 10) *Juegos, diversiones y entretenimiento*; 11) *Deportes: prendas de vestir y objetos que se utilizan*; 12) *Acciones que normalmente realizamos todos los días*; 13) *La naturaleza, las vacaciones*; 14) *Agricultura y vegetales, ganadería y animales, pesca y peces*; 15) *Dinero, banco, comercios*.

disponibilidad y el contenido de manuales de español como segunda lengua o lengua extranjera utilizados en España y en el extranjero y destaca que en ellos se encuentran palabras con un alto índice de disponibilidad junto con otras con una disponibilidad muy baja (Benítez Pérez, 1994b).

Una vez comentados los trabajos pioneros en España, se mencionan las principales investigaciones en las diecisiete comunidades autónomas y las dos ciudades autónomas por orden alfabético. En Andalucía, el proyecto de disponibilidad léxica está dirigido por López Morales y García Marcos (1995). Actualmente se han recopilado los léxicos disponibles de todas las provincias: en Almería realizado por Mateo García (1998), en Cádiz por González Martínez (2002), en Córdoba por Bellón Fernández (2003)⁸⁴, en Huelva por Prado Aragonés y Galloso Camacho (2005), en Málaga por Ávila Muñoz (2006), en Jaén por Ahumada Lara (2006), en Granada por Pastor Millán y Sánchez García (2008) y en Sevilla por Trigo Ibáñez (2011).

Existen numerosos estudios que complementan a los anteriores y que en algunos casos incorporan otros centros de interés o se basan en una muestra de estudio diferente. Mateo García (1994) compara el léxico del centro de interés *Alimentación* de una muestra de Motril con el léxico disponible de Madrid recopilado por Benítez Pérez (1992); Cruzado Romero (2003) ha realizado el estudio sobre los preuniversitarios de Moguer y Ortolano Ríos (2005) sobre Ayamonte; García Megía (2003) estudia el léxico disponible de niños almerienses entre 9 y 12 años; Padro Aragonés y Galloso Camacho analizan el léxico de los escolares de sexto de Primaria de Huelva; Ávila Muñoz y Villena Ponsoda (2010, 2011) analizan el léxico disponible de Málaga sobre una muestra preestratificada de adultos, etc.

Asimismo, habría que incluir la investigación realizada en Gibraltar, territorio de ultramar de Reino Unido ubicado en la región andaluza, donde Escoriza Morera (2003, 2007) analiza el léxico disponible español en contacto con el inglés. En español como lengua extranjera destaca la tesis doctoral de Sánchez-Saus Laserna

⁸⁴ En el año 2003 realizó su tesis doctoral. No obstante, se ha consultado la obra publicada como fruto de la tesis en el año 2011, *Léxico disponible de la provincia de Córdoba*.

(2011) sobre el léxico disponible de estudiantes extranjeros que realizan una estancia en Andalucía.

En Aragón, el equipo de investigación formado por Castañer Martín, Enguita Utrilla, Lagüéns Gracia, Beatriz Moliné y coordinado por Arnal Purroy publica en el año 2004 el léxico disponible de esa comunidad. La muestra está formada por informantes preuniversitarios de las provincias de Huesca, Teruel y Zaragoza. En el año 2008, el grupo de investigadores presenta un estudio desde diferentes disciplinas (sociolingüística, psicolingüística, lingüística aplicada a la enseñanza de lenguas, etc.) a partir de los datos recopilados anteriormente. Paralelamente, Arnal Purroy (2008b) analiza la incidencia de los niveles socioculturales en el léxico disponible de Aragón y en el año 2009 estudia los dialectalismos presentes en los listados.

En Asturias, se publica en el año 2001 el léxico disponible de esa comunidad gracias a la labor de Carcedo González. Además de las variables sociológicas más frecuentes como son el sexo, el centro escolar, el tipo de comunicad y el nivel sociocultural evalúa la zona geográfica desde un punto de vista lingüístico. Para ello, Carcedo González divide la comunidad en cuatro zonas según la variante lingüística: zona gallego asturiano, zona bable⁸⁵ central, zona bable occidental y zona bable oriental. Interesado en las variantes lingüísticas, el propio autor lleva a cabo con posterioridad estudios comparativos con otras sintopías peninsulares e hispanoamericanas (Carcedo González, 2003 y 2004).

En Cantabria, Fernández Juncal (2013) recopila el léxico disponible proveniente de once municipios de las localidades de Santander, Torrelavega, Catro-Urdiales, Mueridas, Santoña, Reinosa, San Vicente, Ampuero, Castañeda, Meruelo y Potes. Con anterioridad, Carrera de la Red y Bradley (2004a y 2004b) presentan los estudios de disponibilidad en Cantabria, llevan a cabo un análisis de los extranjerismos e inician la recopilación de datos que Fernández Juncal continúa y edita. Borrego Nieto y Fernández Juncal (2002) analizan el léxico de los habitantes de zonas rurales y se centran en el estudio de los dialectalismos. García Casero (2013) ha estudiado el léxico disponible en Santander de estudiantes de último curso de la Enseñanza Secundaria Obligatoria.

⁸⁵ El bable hace referencia a la lengua tradicional de la mayor parte del Principado de Asturias.

En Castilla la Mancha, Hernández Muñoz (2006) presenta el léxico disponible de la comunidad. La muestra está formada por informantes de las cinco provincias castellano-manchegas: Ciudad Real con 220 sujetos, Toledo con 180, Albacete con 120, Guadalajara con 80 y Cuenca con 80. Cabe destacar que anteriormente Hernández Muñoz (2004) había estudiado el léxico disponible de Cuenca sobre una muestra de 117 informantes.

En Castilla y León, los resultados del léxico disponible no están disponibles en todas las provincias. Galloso Camacho (2002a, 2002b y 2003) recoge los datos de Salamanca, Ávila y Zamora; Bartol Hernández (2004) de Soria; Fernández Juncal (2008) de Burgos y de Santiago Guervós (2008) de Segovia. Además, esta comunidad destaca por coordinar el programa Dispolex y por publicar una gran variedad de artículos: sobre los dialectalismos y el nivel de instrucción (Fernández Juncal y Borrego Nieto, 2002, 2003); el concepto de disponibilidad (Borrego Nieto, 2004); los extranjerismos (Bartol Hernández, 2005); los regionalismos de Ávila (Llorente Pinto, 2005); sobre el léxico disponible de inmigrantes en la comunidad (Fernández-Merino Gutiérrez, 2015), etc. Actualmente, González está realizando el léxico disponible de Palencia y Valladolid con lo que solo faltaría la provincia de León para completar los estudios de disponibilidad en esa comunidad.

En Cataluña, Serrano Zapata (2003) realiza un estudio exploratorio sobre el léxico disponible de Lérida con una muestra de 48 estudiantes preuniversitarios de tres centros educativos de la capital. A continuación, estudia más en profundidad la influencia de factores sociales (2004) y de transferencias léxicas (2006). En el año 2014, presenta su tesis doctoral sobre el léxico disponible de Lérida sobre una muestra de 240 sujetos preuniversitarios. Los informantes realizan las encuestas en español y catalán sobre 17 centros de interés (16 centros tradicionales y *Los colores*) contribuyendo así a los estudios realizados en comunidades bilingües como el País Vasco (Azurmendi Ayerbe, 1983).

Blasco, Torres, Pérez Edo y Forment han recogido los datos de Barcelona y presentan los aspectos metodológicos del estudio en el *XXXIII Simposio de la Sociedad de Lingüística* celebrado en Gerona en diciembre de 2003. Asimismo, Forment y Pérez

han recopilado los datos de Tarragona y Gerona durante el año 2005 y 2006 pero no todavía no han publicado los listados del léxico disponible.

En Ceuta, la investigadora Ayora Esteban (2006) analiza el léxico disponible de Ceuta con una muestra de informantes de los cinco centros educativos de la ciudad. Entre las variables socioculturales se incluyen el lugar de nacimiento del informante y de los padres y la lengua materna ya que algunos estudiantes son de procedencia marroquí (en la etapa postobligatoria la presencia es menor que las etapas obligatorias). Este estudio forma parte del grupo de investigación Semaínein que pretende desarrollar el léxico disponible de Ceuta, Melilla y Gibraltar y que analiza el contacto de lenguas. Anteriormente, Ayora Esteban estudia centra dos de sus publicaciones en ese aspecto, «La disponibilidad léxica en una situación de contacto de lenguas» en 2002 y «Deficiencias en la adquisición del léxico como problema específico de las lenguas en contacto» en 2004. En obras posteriores, la autora destaca los objetivos y aplicabilidad de los estudios de disponibilidad léxica (Ayora Esteban, 2005 y 2007).

En Extremadura, el léxico disponible está en fase de elaboración. Prado Aragonés y Galloso Camacho se han ocupado de la recogida de datos de 700 estudiantes de las provincias de Cáceres y Badajoz. Asimismo, las autoras recopilaron el léxico disponible de Huelva de estudiantes preuniversitarios (2005) y de 6.º de Educación Primaria (2008). En el año 2009 junto con Conceição presentaron el léxico disponible de zonas limítrofes de Andalucía y Extremadura y de Algarve y Alentejo. Las provincias seleccionadas por el estudio fueron: Oliva de la Frontera, Alconchel, Olivenza, Alburquerque y San Vicente de Alcántara en Badajoz; Valencia de Alcántara, Alcántara y Moraleja en Cáceres; Ayamonte, Puebla de Guzmán y Cortegana en Huelva; Vila Real en Algarve; Mértola, Serpa, Moura, Vila Viçosa, Reguengos de Monsaraz, Elvas y Portalegre en la Región de Alentejo (Portugal). La metodología de investigación sigue las recomendaciones del Panhispánico y la encuesta se realizó a una muestra de 240 estudiantes preuniversitarios españoles y 160 portugueses. El promedio de respuestas de los informantes españoles es superior al de los estudiantes portugueses pero el porcentaje de vocablos es inferior.

En la Comunidad Valenciana, se ha elaborado el léxico de Valencia (Gómez Molina y Gómez Devís, 2004) y Alicante (Martínez Olmos, 2008⁸⁶) mientras que el léxico de Castellón se encuentra en fase de elaboración por parte de Blas y Casanova Ávalos. Esta última investigadora, presenta los primeros resultados de su estudio en 2006 y analiza los anglicismos que aparecen en el listado del léxico de Castellón en el año 2008. Asimismo, Llopis Rodrigo y Gómez Devís (2010a, 2010b) siguen las recomendaciones del Panhispánico para estudiar la disponibilidad léxica del catalán en Valencia.

En Galicia, López Meirama (2008) presenta el léxico disponible de la comunidad sobre una muestra de 800 informantes. Para la recogida de datos no estableció la distinción según la demarcación provincial por considerar la comunidad como un conjunto. Para ello, modifica la variable relativa al tipo de población y distingue entre urbana, periurbana, villa y rural. Además de los centros tradicionales incluye *Los colores* y *El mar*.

Esta investigación se inscribe dentro del proyecto *Disponibilidad léxica en Galicia* cuyo objetivo principal es la realización de diccionarios de léxico disponible en español y en gallego, además de crear una herramienta informática para su consulta. Los miembros que forman parte del proyecto son López Meirama, Álvarez de la Granja, Blanco y García Gondar. En el año 2011 publican una obra conjunta, en la que participa Bartol Hernández, donde se abordan las diferencias dialectales del español y del gallego, la ortografía, la incidencia del tipo de población, entre otros aspectos desde una perspectiva multidisciplinar.

En las Islas Canarias, se ha recopilado el léxico disponible de Las Palmas de Gran Canaria y Tenerife. García Domínguez, Marrero Pulido, Pérez Martín y Piñero Piñero (1994) estudian el léxico disponible de Gran Canaria y analizan la incidencia de la situación geográfica y el tipo de educación. Previamente, Samper Padilla y Hernández Cabrera (1997) presentan los datos iniciales del estudio de disponibilidad léxica en Gran Canaria y dan comienzo a una serie de investigaciones en las islas.

En 1998, Samper Padilla publica un artículo sobre los criterios de edición del material que serán tenidos en cuenta en la mayoría de investigaciones posteriores en

⁸⁶ Tesis doctoral presentada en 2007.

este ámbito. Hernández Cabrera, Samper Padilla y Bellón Fernández (2003) analizan los canarismos en los listados de disponibilidad léxica y Hernández Cabrera y Samper Padilla (2006) abordan los dialectalismos. Asimismo, el léxico gran canario ha formado parte de comparaciones con otros léxico, tales como el de Puerto Rico (Samper Padilla, 1999), Córdoba (Samper Padilla, Hernández Cabrera y Bellón Fernández, 2003) y Aragón (Samper Hernández, 2008), entre otros.

En Tenerife, Torres González y García Falero muestran algunos datos sobre el léxico de la isla en el *III Congreso Internacional sobre Lengua y Sociedad* celebrado en Castellón en el año 2006. En el resto de islas, Samper Hernández se encarga de la recopilación de datos. Para su tesis doctoral, la investigadora realiza un estudio comparativo entre los escolares grancanarios de primaria y secundaria para analizar su evolución (Samper Hernández, 2009) y sobre el que previamente había aportado datos metodológicos (Samper Hernández, 2004) y analizado los dialectalismos (Samper Hernández, 2005).

En Madrid, nace el primer estudio que sigue los criterios del Panhispánico a cargo de Benítez Pérez (1992)⁸⁷. En Alcalá de Henares, Paredes García, Álvaro García, Pérez, Rebollo y Ruiz (2002) estudian el léxico disponible de los estudiantes del segundo ciclo de Educación Secundaria Obligatoria (3.º y 4.º) y Bachillerato y lo comparan con el obtenido por Benítez Pérez (1992). Además, Paredes García analiza aspectos concretos del léxico como los colores (2006), los errores ortográficos (2011) y los avances teóricos y metodológicos en el ámbito de la disponibilidad léxica (2012).

El léxico de la comunidad madrileña está en fase de elaboración por el equipo de investigación compuesto por Paredes García (Universidad de Alcalá) y Gómez Sánchez y Guerras Salas (Universidad Europea de Madrid). En la comunidad, existen estudios que analizan el español en informantes extranjeros como el de Guerra Salas y Gómez Sánchez (2003, 2005a, 2005b) con una muestra de estudiantes extranjeros de Comunicación de la Universidad Europea de Madrid y la tesis inédita de Diego Gallego Gallego (2014) sobre estudiantes de español como lengua extranjera en Madrid, entre otros.

⁸⁷ Guerra Salas, Paredes García, y Gómez Sánchez presentan la comunicación titulada «Disponibilidad léxica en la Comunidad de Madrid» en el «XXXIV Simposio de la Sociedad Española de Lingüística» celebrado en Madrid del 13 al 16 de diciembre de 2004.

En Melilla, el léxico disponible forma parte del proyecto de estudio del léxico en Ceuta, Melilla y Gibraltar llevado a cabo por el grupo Semaínein. Fernández Smith, Rico Martín, Molina García y Jiménez Jiménez presentaron el proyecto en el año 2006 y dos años más tarde, en 2008, publicaron el léxico disponible de Melilla. Al igual que en Ceuta, utilizaron la variable lengua materna pero en este caso diferenciaron entre español, bereber y otras.

En Navarra, el léxico disponible no ha sido publicado todavía pero el equipo formado por Tabernero Sala⁸⁸, Saralegui, Martínez, Areta Lara, Mangado y Sanvicens de la Universidad de Navarra han recopilado los datos. En el año 2007, Saralegui y Tabernero Sala (2008⁸⁹) participaron en el XXXVII *Simposio Internacional de la Sociedad Española de Lingüística* (SEL) con la comunicación relativa a la aportación al Panhispánico del léxico disponible de Navarra.

Ese mismo año, Sanvicens Gila (2007) realiza su trabajo de investigación de Tercer Ciclo donde analiza los aspectos teóricos del léxico disponible en comunidades bilingües y presenta el caso de Navarra. Un año más tarde, Areta Lara (2008⁹⁰) presenta su tesis doctoral sobre el léxico disponible de estudiantes preuniversitarios navarros y realiza un análisis sociolingüístico pormenorizado. En el ámbito del español para extranjeros, Jiménez Berrio (2013) analiza la competencia léxica de inmigrantes en Secundaria, relaciona la disponibilidad léxica con el aprendizaje integrado de contenidos y lenguas (Jiménez Berrio, 2012a) y los programas de inmersión lingüística (Jiménez Berrio, 2012b), entre otros.

El País Vasco es la primera comunidad en presentar un estudio relacionado con el léxico disponible gracias a la tesis doctoral de Azurmendi Ayerbe (1983). En el año 1996, Etxebarria Arostegui realiza un estudio sobre el léxico de los estudiantes preuniversitarios de Bilbao del que más tarde, en 1999, analiza la enseñanza de la lengua materna. La investigación se complementa con los estudios que comienzan en 2005 en las provincias de Álaba, Vizcaya y Guipúzcoa y que están todavía en curso.

⁸⁸ Ese mismo año la autora publica un artículo relacionado con los regionalismos (Tabernero Sala, 2008a) y el contacto de lenguas (Tabernero Sala, 2008b).

⁸⁹ Las actas del congreso se publican un año más tarde.

⁹⁰ Mangado y Areta Lara (2008) publican en las actas del «XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)» un artículo relacionado con el procesamiento informático de los datos.

2.4.2. En inglés como lengua extranjera

En el ámbito del inglés como lengua extranjera las investigaciones son menos numerosas. El yugoslavo Dimitrijević realiza el primer estudio en inglés en 1969. Parte de hipótesis que el léxico disponible puede verse influido por variables como la edad, el sexo, el estatus social y el país (Dimitrijević, 1969: 5). Para corroborar esa teoría, recopila datos en dos países: Escocia y Yugoslavia. En el primer caso, selecciona a niños que viven en zonas rurales y urbanas de Escocia de diferentes edades mientras que, en el segundo, selecciona una muestra de escolares de diferente edad (11 y 14 años) monolingües serbo-croatas y bilingües húngaros con el serbo-croata como segunda lengua. Sin embargo, en la obra publicada en 1969 solo compara los resultados de los escolares escoceses.

No es hasta el año 2000, cuando se publican los resultados de una investigación realizada por Germany y Cartes Enríquez sobre el léxico disponible en inglés como lengua extranjera en la enseñanza formal de la Octava Región de Chile. La encuesta la realizan 100 escolares de primer año de enseñanza media sobre tres centros de interés: 1) *Human body*; 2) *Food*; 3) *House*. El objetivo consiste en determinar el léxico de los estudiantes según la variable tipo de centro educativo: 28 informantes estudian un establecimiento escolar bilingüe, 38 en un establecimiento escolar particular y 40 en un liceo municipalizado.

Tras realizar el análisis de los resultados, Germany y Cartes Enríquez (2000: 49) constatan que los escolares del colegio bilingüe aportan un mayor número de palabras y de vocablos que sus compañeros en los otros dos tipos de centros. A continuación, le siguen por poca diferencia los resultados del colegio particular y, por último, los del liceo municipalizado donde se encuentran con encuestas sin ninguna respuesta. Según las autoras, los resultados de los dos primeros centros educativos se justifican por la metodología aplicada en la enseñanza del inglés: mediante semi-inmersión en el colegio bilingüe y siguiendo un método comunicativo en el colegio particular. En cambio, en el liceo municipalizado la enseñanza del inglés se lleva a cabo durante solo tres horas semanales basando la metodología en el aprendizaje gramatical.

En el año 2006, dos estudiantes de la Universidad de Concepción, Ferreira y Zubanov, presentan sus tesis del Magíster en Lingüística relacionadas con la disponibilidad léxica en lengua extranjera. El trabajo de Ferreira (2006) analiza las diferencias entre el léxico disponible en inglés como lengua materna y como lengua extranjera a través de un análisis cuantitativo y psicolingüístico. Los datos de las pruebas provienen de 50 hablantes nativos de entre 16 y 18 años del sur de Inglaterra inscritos en The Royal School y 50 alumnos chilenos de cuarto curso de las carreras de Pedagogía en Inglés y de Traducción en Idiomas Extranjeros de entre 21 y 22 años.

Entre los centros seleccionados, los cuatro primeros centros corresponden a los centros tradicionales y los seis restantes son específicos del estudio: 1) *Body parts*; 2) *Food and drink*; 3) *Entertainment*; 4) *Clothes*; 5) *Holidays*; 6) *Economy and finances*; 7) *Terrorism and crime*; 8) *Politics*; 9) *Pollution and the environment*; 10) *Health and medicine*. Los resultados presentados muestran que el número de palabras y vocablos en todos los centros es superior en los hablantes de inglés como lengua materna. Además, el vocabulario proporcionado por los escolares ingleses es más especializado y específico mientras que el léxico de los estudiantes chilenos es más homogéneo, debido a que la exposición a la lengua extranjera se produce en un ambiente muy similar, el escolar.

A través de un análisis realizado con el programa DispoGrafo, el investigador presenta la organización de los vocablos en los dos listados. De esta forma, se observa que los hablantes nativos organizan las palabras en el lexicón mental a través de subnúcleos en cada centro de interés que revelan la relación semántica de las palabras. En cambio, esa relación en subcategorías no es tan evidente en el caso de los estudiantes chilenos ya que las palabras se integran en un centro sin una clara organización dejando patente que su dominio léxico es solo parcial.

El estudio de Zubanov (2006) se basa en el progreso léxico de estudiantes durante sus estudios universitarios en inglés y en alemán. En inglés como lengua extranjera, se llevan a cabo encuestas a 80 alumnos divididos en cuatro grupos de diferentes cursos de la carrera de Traducción en Idiomas Extranjeros de la Universidad de Concepción. Para poder comparar el léxico disponible cuantitativa y

cualitativamente con hablantes de lengua materna, la encuesta la realiza del mismo modo a un grupo de 20 estudiantes norteamericanos.

Como era de esperar, los hablantes nativos muestran un vocabulario más amplio y rico que los estudiantes de inglés como lengua extranjera. La autora destaca algunas diferencias culturales, tales como, la mayor asociación de palabras positivas en el centro de los estados de ánimo en los estudiantes chilenos y un mayor número de palabras afectivas en el centro de los animales. En cuanto a la evolución del léxico en los estudiantes chilenos, destaca que se produce de forma progresiva a medida que los estudiantes adquieren la lengua extranjera.

El estudio exploratorio llevado a cabo por Gallardo del Puerto y Martínez Adrián (2014) tiene por objeto conocer la influencia del contacto de una lengua extranjera en el léxico disponible de estudiantes de la tercera edad. Los datos provienen de 18 informantes de edad avanzada que estudian inglés como lengua extranjera en las Aulas de Experiencia de la Universidad del País Vasco durante tres horas semanales. Los encuestados se dividen, según su experiencia previa con el inglés como lengua extranjera, en dos grupos: uno formado por estudiantes principiantes que parten de cero denominados *True Beginners (TBs)* y otro grupo compuesto por los falsos principiantes⁹¹ o *False Beginners (FBs)* que son estudiantes que comienzan a estudiar una lengua desde cero pero tienen un ligero conocimiento de la misma.

La prueba léxica se realiza a partir de quince centros de interés relacionados con la vida diaria: 1) *Parts of the body*; 2) *Clothes*; (3) *House*; (4) *Furniture*; (5) *Food and drink*; 6) *Table*; 7) *Kitchen*; 8) *School*; 9) *Town*; 10) *Countryside*; 11) *Means of transport*; 12) *Animals*; 13) *Hobbies*; 14) *Professions*; 15) *Colours*. Los tres factores objeto de análisis son: el número de palabras en general, la determinación de los centros más productivos en ambos grupos y los vocablos aportados en cada centro de interés. El número de palabras corrobora la hipótesis de los investigadores ya que muestran que

⁹¹ La definición de este concepto puede consultarse en la versión electrónica del *Diccionario de términos clave de ELE* del Instituto Cervantes. Disponible en: http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/falsoprincipiante.htm [Consultado el 9 de marzo de 2015].

los falsos principiantes superan de manera significativa a los principiantes en el número total de palabras producidas en la mayoría de los centros de interés.

La productividad en los diferentes centros es muy parecida en los dos grupos del estudio y en otras muestras con informantes de diferente edad que analizan la disponibilidad en lengua extranjera (Germany y Cartes 2000; Jiménez Catalán y Ojeda Alba 2009, 2010). Los centros más productivos son *Food and drink*, *Animals* y *Town* mientras que *Kitchen* y *Table* son los menos productivos. Las semejanzas observadas en los estudiantes de edad avanzada y otras poblaciones, permiten sugerir que los estudiantes de nivel inicial parecen experimentar un desarrollo similar para adquirir vocabulario con independencia de la edad. Asimismo, el análisis cualitativo de las palabras más disponibles en cada centro de interés, muestra una gran coincidencia entre los dos grupos. Por último, los autores concluyen que la habilidad de aprender una lengua en los adultos no se ve disminuida gracias a su alta motivación por aprender y, por tanto, los factores de la edad no tienen tanto impacto como se suele esperar. Además, proponen realizar estudios más exhaustivos sobre el impacto de ciertas variables relacionadas con la edad, como la pérdida de la capacidad auditiva y visual en el proceso de adquisición del léxico de una lengua extranjera.

Con motivo del *V Simposio para el mejoramiento de la enseñanza y el aprendizaje de las lenguas, una necesidad formativa urgente de atender para el logro de la educación omnilateral en el siglo XXI* celebrado en mayo de 2014 en México, Carlos Fuentes Páez (2014) presenta las experiencias iniciales de un estudio de disponibilidad léxica⁹² en inglés como lengua extranjera de alumnos mexicanos. La muestra de estudio está formada por 59 escolares de cuarto año de preparatoria (desde 14 años) de los cuales 34 cursan sus estudios en el sector público y 25 en el sector privado en la zona norte del Distrito Federal de la Ciudad de México.

La prueba de léxico disponible se realiza en dos periodos, a principios del curso 2012 y tras la evaluación bimestral, sobre cuatro centros de interés: 1) *Countries and nationalities*; 2) *Adjectives*; 3) *Family members*; 4) *Jobs and occupations*. Durante el tiempo transcurrido entre una y otra prueba, los alumnos realizan diferentes tipos de

⁹² En http://cel.prepa2.unam.mx/documentos/eventos/5osimposio/M2_14.pdf [Fecha de última consulta el 20 de marzo de 2015.]

ejercicios para reforzar ese léxico. El objetivo del estudio es conocer el número de palabras que escriben los alumnos de uno y otro sector en dos momentos diferentes para comprobar la progresión del aprendizaje de los mismos.

Los resultados del estudio muestran que los escolares pertenecientes al sector privado escriben más unidades léxicas que los estudiantes del sector público en todos los centros. Los alumnos mostraron una progresión léxica en todas las áreas de interés, especialmente en *Países y nacionalidades* y, en menor medida, en el centro de adjetivos ya que presentaban muchos errores ortográficos y esas palabras no se computan en los resultados. No obstante, los estudiantes del sector público muestran una mejor progresión entre una y otra encuesta, acortando la diferencia numérica con respecto a las palabras del sector privado en los centros de interés *Miembros de la familia* y *Trabajos y ocupaciones*. El autor concluye que el tipo de ejercicios realizados para fomentar la competencia léxica ha sido bastante fructífero, con independencia del sector en el que se aplique, llegando a disminuir las diferencias entre el sector público y privado.

2.5. Revisión del modelo matemático

2.5.1. Cálculo del índice de disponibilidad léxica

El vocabulario recogido en el *Français Fondamental* consta de dos pruebas complementarias⁹³. La primera, el léxico frecuente, consiste en recopilar las palabras presentes en diferentes grabaciones con referencia al número de menciones y de intervenciones en los que aparece⁹⁴. Concretamente, los investigadores franceses recopilaron 301 grabaciones de las cuales eliminaron 31 por la escasez de su discurso. De los 275 hablantes restantes, 138 era hombres, 126 mujeres y 11 niños en edad escolar. Además, analizaron algunas características de los hablantes tales como el reparto por región, el tipo de profesión y el empleo.

El resultado era una lista de palabras en las que se incluía el número de frecuencia o número de apariciones y el número de documentos sonoros en los que

⁹³ Vid. apartado 2.2 relativo a los orígenes de la disponibilidad léxica.

⁹⁴ En investigaciones anteriores, ese léxico frecuente se extraía a partir de documentos escritos de diversa índole (Henmon, 1924; Vander Beke, 1929; Ogden, 1930; Tharp, 1939; entre otros).

aparecía (*répartition*⁹⁵). Asimismo, para evitar ambigüedades en algunos casos se incluía la referencia gramatical de la palabra (verbo, adverbio, pronombre, preposición, etc.). Gougenheim *et al.* (1956 y 1964) presentaron los listados por orden decreciente de frecuencia y por orden alfabético.

Liste des mots par fréquences décroissantes			
Numéros d'ordre	Mots	Répartition	Fréquence
1	être (verbe)	163	14.083
2	avoir	163	11.552
3	de	163	10.503
4	je	162	7.905
5	il(s)	160	7.515
6	ce (pronom)	163	6.846
7	la (article)	163	5.374
8	pas (négation)	158	5.308
9	à (préposition)	163	5.236
10	et	161	5.082

Tabla 2.1. Ejemplo del listado de palabras clasificadas según orden de frecuencia. Extraído de Gougenheim *et al.* (1964: 69)

En cuanto al listado de palabras disponibles, solo se tenía en cuenta el número de menciones, es decir, el número de frecuencias. En este caso, no se planteaba la posibilidad que un informante escribiese dos veces la misma palabra, con lo que el número de menciones coincidía con el número de informantes que habían participado en la prueba. La Tabla 2.2 recoge un ejemplo de los 30 vocablos más mencionados en el centro de interés relativo a la ropa. Los datos pertenecen a la prueba realizada en el departamento de Dordogne y está formada por 4 grupos: 120 chicos de medio urbano, 120 chicas de medio urbano, 130 de medio rural y 130 chicas de medio rural. En total 250 chicos y 250 chicas, 240 de ciudad y 260 del campo. Esa misma lista podría presentarse filtrando los datos según las variables sexo y medio ya que los autores franceses realizan algunos análisis comparativos teniendo en cuenta esas dos variables.

Dordogne. Les vêtements					
N° d'ordre	Mots	Nombre de mentions	N° d'ordre	Mots	Nombre de mentions
1	chemise	472	16	impermeable	202
2	veste	442	17	blouse	198

⁹⁵ Si aparecía en menos de 20 conversaciones ese vocabulario era eliminado con independencia de su frecuencia.

3	tablier	433	18	chapeau	198
4	manteau	413	19	corsage	192
5	gilet	408	20	pull-over	189
6	culotte	389	21	cache-nez	179
7	robe	382	22	blouson	171
8	chaussettes	376	23	caleçon	169
9	jupe	344	24	corset	163
10	combinaison	324	25	slip	132
11	pantalón	266	26	socquette	131
12	cravate	254	27	souliers	115
13	bas	217	28	canadienne	104
14	gants	214	29	soutien-gorge	98
15	béret	205	30	pardessus	95

Tabla 2.2. Ejemplo del listado de vocablos ordenados según el número de menciones. Extraído de Gougenheim *et al.* (1964: 157)

La metodología seguida para el recuento de palabras estará presente en estudios ulteriores en diferentes partes del mundo, tales como en Canadá (William Mackey, 1971a, 1971b), en Inglaterra (Dimitrijević, 1969), en Puerto Rico (López Morales, 1973, 1978, 1979; Román, 1985) y en España (Azurmendi Ayerbe, 1983). A pesar de que Michéa (1953), Gougenheim *et al.* (1956) y Müller (1969) advirtieron de la necesidad de incluir el orden de aparición de los vocablos en algún tipo de índice o recuento, no será hasta 1978 cuando Njock establezca la distinción entre el índice de frecuencia y el rango asociativo de la palabra.

Para calcular el rango asociativo de la palabra, Njock se basa en el concepto de rango individual que corresponde a la posición que ocupa esa palabra en la lista escrita por los informantes⁹⁶. Ese rango individual es diferente al rango del concepto, el cual hace referencia al puesto que ocupaba la palabra según el orden decreciente de frecuencia léxica.

Il faut introduire une notion supplémentaire, celle de rang, et classer les vocables par fréquence décroissante, en inversant les tableaux de distribution tels que nous les avons présentés; on attribue alors le rang 1 au vocable le plus fréquent et ainsi de suite; pour les vocables de même fréquence, on prend un rang moyen; sous sa forme la plus simple (Müller, 1968: 167).

⁹⁶ «Le rang individuel du mot est celui que le concept a dans la liste des réponses d'un élève. Si un sujet trouve 50 mots au centre d'intérêt 1 (les parties du corps humain) et que les concepts tête et appendice aient respectivement été fournis en premier et dernier lieu, nous dirons que tête a le rang individuel 1, et appendice le rang individuel 50» (Njock, 1978: 119).

Tras exponer la importancia asociativa de las palabras, Njock (1978: 201) presenta la fórmula para calcular lo que determina como rango asociativo (R_a). En primer lugar, calcula el rango medio (R_m) del concepto en el centro de interés añadiendo los diferentes rangos individuales que el concepto ha obtenido ($r_a + r_b + \dots + r_n$), después divide el total ($\sum r_a - n$) por el número de sujetos (n) que han escrito ese vocablo.

$$R_m = \frac{\sum r_a - n}{n} \text{ o } r_a < R_m < r_n$$

A continuación, toma como referencia el número de sujetos (X) que ha escrito el concepto con un resultado igual o inferior a la media calculada (R_m). Por último, calcula el porcentaje del total (X) obtenido con respecto al número de sujetos (N) que han respondido a la encuesta⁹⁷.

$$R_a = \frac{x \cdot 100}{N} = \frac{x}{2,2}$$

Seguidamente, el investigador plantea la cuestión de si realmente el rango asociativo y el índice de disponibilidad miden aspectos diferentes⁹⁸. Para ello, presenta tres ejemplos que ilustran claramente dos índices diferentes (*cf.* Njock, 1978: 201-202):

1. El rango asociativo no es proporcional al índice de disponibilidad. Una palabra puede ser muy disponible sin que su rango asociativo sea tan importante en comparación con el de la palabra disponible. Por ejemplo, la palabra *tarot* en francés es más fácil de escribir que *libxanga* en basaa, la disponibilidad en ambas lenguas es muy parecida 58,1 en francés y 59,5 en basaa. Sin embargo, en francés el rango es superior 42,2 que en basaa 37,2 ya que suele aparecer más tarde.
2. El rango asociativo parece medir, en cierto modo, la importancia de la palabra en la lengua del sujeto. En palabras que son aceptadas de la misma forma en los dos idiomas, el rango asociativo será el mismo. Es el caso de la mayoría de las palabras en las que el locutor no hace gran diferencia entre el francés y el basaa y, por ende, no percibe ciertas palabras como préstamos.

⁹⁷ En su estudio $N = 220$ de ahí que en la fórmula ya se incluya la cifra 2,2.

⁹⁸ «Une question que nous nous posons est de savoir ce que mesure le rang associatif? N'est-il pas la même chose que l'indice de disponibilité? Peut-il nous aider à mieux apprécier l'indice de disponibilité?» (Njock, 1978: 201).

Para esas palabras tan aceptadas, el rango también es muy importante en las dos lenguas. Por ejemplo, *tennis* tiene un índice de disponibilidad en basaa de 4,5 mientras que en francés 8,1 y, sin embargo, el rango asociativo en ambos casos es de 2,7.

3. Las palabras con un mismo índice de disponibilidad no tienen siempre un rango asociativo similar. Por tanto, el rango asociativo no mide lo mismo que el índice de disponibilidad. La Tabla 2.3 muestra palabras con un mismo índice de disponibilidad pero diferente rango asociativo.

Rang associatif pour des concepts de même indice de disponibilité											
Rang	Basaa	Disponibilité	Rang Associatif	Rang	Français	Disponibilité	Rang Associatif	Rang	Basaa et Français	Disponibilité	Rang Associatif
18	hyonx (cheveu)	43,6	23,6	2	oeil	87,2	57,2	35	koxox (peau)	17,2	8,6
18	jas (menton)	43,6	22,2	2	tête	87,2	67,2	35	crâne	17,2	8,6
26	bxep (lèvre)	22,2	10,9	11	dent	66,8	37,7	14	tol (poitrine)	58,1	32,7
26	nla (intestin)	22,2	12,7	11	pied	66,8	36,3	14	genou	58,1	29,5
28	nxexm (coeur)	21,8	10,4	17	main	51,8	30,9	15	limanx (joue)	51,8	29,5
28	hes (os)	21,8	12,2	17	joue	51,8	29,5	17	joue	51,8	29,5
30	tindi (talon)	21,3	11,8	27	orteil	32,2	17,7	17	main	51,8	30,9
30	lbxee (mamelle)	21,3	11,3	27	tronc	32,2	19,5	33	sourcil	18,1	7,2
40	su (figure)	11,3	7,2	35	cœur	17,2	8,6	34	tuu (épaule)	18,1	10,4
40	yijjis (sourcil)	11,3	4,5	35	crâne	17,2	8,6	38	libxit (gorge)	11,8	7,2
								42	foie	11,8	5,4

Tabla 2.3. Listado de palabras con un mismo índice de disponibilidad pero diferente rango asociativo. Centro de interés 01. *Parties du corps humain*. Extraído de Njock (1978: 203)

En esta misma línea, se desarrollan otras fórmulas que incorporan la posición que ocupan las palabras en los listados. Los trabajos de Lorán Santos (1983) y Lorán Santos y López Morales (1983) ponderan la frecuencia de las palabras según su orden de mención. Según Lorán (1987), la fórmula inicial se basa en la teoría de las decisiones y consiste, *grosso modo*, en multiplicar un coeficiente con valor inferior a la unidad (λ) por la frecuencia de la unidad. Ese cálculo no era efectivo en algunos casos. Por ejemplo, cuando las palabras aparecían en primer lugar el coeficiente no actuaba y, por tanto, la frecuencia era la misma. Tras varios intentos de mejorar el cálculo de la disponibilidad de una palabra, los investigadores propusieron una fórmula para los listados de palabras con un tamaño homogéneo que alteraba la tasa de sustitución si el tamaño de los listados era diferente.

$$\text{Por un lado: } d(p) = \sum_{i=1}^n x_{pi} \text{ si } x_{pi} = \frac{f_{pi}}{N_i}$$

$$\text{Por otro: } d(p) = \sum_{i=1}^n \lambda^{i-1} \frac{f_{pi}}{N_i}$$

donde

n = máxima posición alcanzada por las palabras en la encuesta

i = número de posición

x_{pi} = frecuencia relativa de la palabra p en la posición i

f_{pi} = frecuencia absoluta de la palabra p en la posición i

N_i = número de informantes cuyas listas contenían la posición i

La disponibilidad se calcula sumando el número de veces que la palabra ocupa cada posición ponderado por una potencia del coeficiente (λ). El valor de (λ) es 0,90 y desciende su potencia conforme disminuye la posición de la palabra en cuestión. Así, el valor del sumando cuando la palabra con la misma frecuencia ocupa la posición 1 será mayor que el del sumando cuando aparezca en cualquier otra posición. El análisis del índice de disponibilidad léxica a partir de la fórmula reseñada está presente en diversos estudios, tales como, Román (1985), Mena Osorio (1986), Justo Hernández (1987), Cañizal Arévalo (1987) y Echeverría *et al.* (1987). Sin embargo, se constata que si la potencia de (λ) es superior a 23, es decir, cuando la palabra se encuentra en la posición 23 o en alguna superior, el sumando deja de ser

significativo. La potencia de (λ) se hace tan pequeña que se produce una curva cercana al eje de abscisas.

Posteriormente, los trabajos presentados por Butrón (1987, 1991) consiguen corregir ese error de forma parcial. La fórmula revisada sustituye el valor (λ) por un número E elevado a un exponente que pondera la posición de la palabra. La investigadora aplica la fórmula en un trabajo realizado en Puerto Rico sobre una muestra de 154 estudiantes de quinto grado donde analiza la influencia de diversas variables, como el sexo, el nivel económico y la procedencia geográfica en la disponibilidad léxica. A pesar de que la fórmula no consigue eliminar por completo el desajuste, el estudio muestra la efectividad de ponderación del número de palabras y la posición de las mismas a través de un número E elevado a un exponente complejo.

En esa misma línea, los profesores de la Universidad Nacional Autónoma de México, López Chávez y Strassburger⁹⁹ consiguen mejorar ese artificio matemático aplicando el número E , conocido como número de Euler o constante de Napier, con un valor 2,718281828459045. La matriz presentada por los estudiosos mexicanos tiene en cuenta los siguientes factores (López Chávez y Strassburger, 1987):

- La frecuencia absoluta con que fue mencionada cada palabra en cada posición.
- La frecuencia absoluta de la palabra, es decir, la suma de las diferentes frecuencias alcanzadas por la palabra en cada posición.
- El número de hablantes que realizan la prueba de disponibilidad léxica.
- El número de posiciones alcanzadas en el centro de interés en cuestión.
- Las posiciones en las que fue mencionada la palabra.

$$D(P_j) = + \sum_{i=1}^n e - C \left(\frac{i-1}{n-1} \right) \frac{f_{ji}}{I1}$$

donde

D hace referencia a la disponibilidad de P_j de una palabra p concreta j

⁹⁹ López Chávez y Strassburger (1987, 1991).

n indica la máxima posición alcanza en el centro de interés

fp_i es la frecuencia absoluta de la palabra en la posición i

$I1$ es el número de informantes que realizaron la prueba

e es el número natural con un valor de 2,718281828459045

La fórmula mexicana es contrastada empíricamente en estudios ulteriores, tales como los realizados en México (López Chávez, 1995), Puerto Rico (López Morales, 1995a y 1995b), Madrid (Benítez Pérez, 1992, 1994a, 1994b), Las Palmas de Gran Canaria (Samper y Hernández, 1994, 1995) y República Dominicana (Alba Ovalle, 1995a y 1995b). En cuanto al tratamiento de los datos, cabe destacar que los propios autores de la fórmula, López Chávez y Strassburger Frías, desarrollaron el programa informático DISPLEX. Esta aplicación permitía no solo calcular el índice de disponibilidad léxica sino también la frecuencia de aparición de las palabras, la frecuencia acumulada, la frecuencia absoluta y la posición que ocupa cada vocablo en un centro de interés con la posibilidad de realizar estratificaciones.

Debido a la complejidad del uso del programa creado en la Universidad Nacional Autónoma de México y su difícil accesibilidad, en 1995 el Grupo de Sociolingüística Teórica y Aplicada de la Universidad de Almería desarrolla el programa ESDIS (Estadística de disponibilidad). En 1998, Mateo García publica su tesis doctoral sobre el léxico de los estudiantes preuniversitarios almerienses en la que procesa los datos mediante este programa.

Para facilitar la aplicación de la fórmula en otros estudios, en 1996 el entonces presidente de la Asociación de Lingüística y Filología de la América Latina (ALFAL), López Morales, encarga a la Universidad de Alcalá la creación de un programa informático: Lexidisp. El programa permite la creación de diccionarios de disponibilidad léxica y el cálculo del índice de disponibilidad, la frecuencia relativa y la frecuencia absoluta. No obstante, el número de variables está restringido a cinco y solo está disponible para el sistema operativo Microsoft Windows.

En el año 2000, se consolida el proyecto salmantino Displex que permite trabajar con una mayor cantidad de datos y variables (Bartol Hernández y Hernández Muñoz, 2004). La principal diferencia en el cálculo de los datos que ofrece Lexidisp y

Dispoplex reside en que el primero es un programa que analiza los datos a partir de ficheros organizados según unas características específicas¹⁰⁰, mientras que el segundo es una base de datos en la que hay que ir introduciendo las referencias manualmente. Además, se podría decir que Dispoplex es un gran banco de datos científico que permite que los investigadores interesados ingresen sus datos, los calculen y los contrasten con los datos de otras investigaciones a las que tengan acceso¹⁰¹.

Echeverría *et al.* (2005) crean una aplicación en MatLab (programa especializado en análisis estadísticos multivariantes) denominada Dispogen¹⁰². Este programa realiza cálculos muy similares a los de Dispoplex (número de vocablos por centro, promedios de respuestas, frecuencia absoluta y relativa, porcentaje de aparición y listado de vocablos por centro), permite conocer los datos de una gran variedad de variables y obtener gráficos con los análisis realizados. Al igual que Lexidisp, se trata de un programa que puede descargarse en cualquier ordenador pero su última actualización es más reciente, concretamente en el año 2008.

En esta misma línea, Frey Pereyra crea un programa denominado *Ledical* que puede ser utilizado en cualquier sistema operativo compatible con Java. La aplicación informática se encuentra disponible en línea y se puede descargar gratuitamente¹⁰³. Este programa es utilizado para procesar los datos de la tesis doctoral de Frey Pereyra (2008), *El léxico disponible en los escritos de alumnos de español como ELE*, así como en otras investigaciones fruto del proyecto de investigación *Estudios de Léxico Disponible* de la Universidad Nacional Autónoma de México a cargo de Marisela Colín Rodea¹⁰⁴.

¹⁰⁰ *Vid.* apartado 3.5.2 referente a la metodología.

¹⁰¹ El acceso a cada investigación es restringido para preservar la privacidad de los datos. No obstante, el investigador responsable de una investigación podrá permitir el acceso a otro u otros investigadores.

¹⁰² La información sobre el proyecto que financia la creación del programa se encuentra disponible en <http://dspace2.conicyt.cl/bitstream/handle/10533/16599/1050598-IF.pdf?sequence=2> [Fecha de última consulta el 5 de marzo de 2015].

¹⁰³ Aunque Frey Pereyra indica que el programa está alojado en la página web www.ledical.org actualmente ese enlace no está disponible. En cambio, se puede acceder al programa previa autenticación en la página <http://cele.unam.mx/ledical/> [Fecha de última consulta el 9 de febrero de 2015].

¹⁰⁴ Para más información sobre el proyecto, consúltese la página <http://dla.cele.unam.mx/terminologia/interiores/proyectos.html> [Fecha de última consulta el 9 de febrero de 2015].

A través del proyecto *Disponibilidad Léxica en Galicia*, financiado por el Programa de Promoción Xeral de Investigación de la Consellería de Innovación e Industria de la Xunta de Galicia, se elaboran dos diccionarios del léxico disponible de los estudiantes preuniversitarios gallegos: uno en español y otro en gallego. Para facilitar la consulta de los datos de las investigaciones inscritas en este proyecto, Guillermo Rojo crea una aplicación informática, Dispogal (López Meirama, 2011: 13). Cabe mencionar que en la propia página del proyecto se indica que los datos han sido procesados con Lexidisp por lo que el programa se utiliza solo como base de consulta¹⁰⁵. No obstante, Dispogal requiere darse de alta como usuario para poder acceder a la información¹⁰⁶.

En el año 2014, Gallego Gallego propone el tratamiento de los datos informáticos a través del programa SPSS. Entre las ventajas de utilizar este programa estadístico para procesar los datos del léxico disponible, Gallego Gallego (2014) destaca la flexibilidad en el uso de diferentes tipos de archivos y en la incorporación de todas las variables que el investigador desee explorar. Asimismo, propone una revisión de la fórmula tradicional propuesta por López Chávez y Strassburger Frías (1987) en la que se aconseja que la constante C tome el valor de 2,3.

2.5.2. Índice de compatibilidad léxica y selección del léxico nuclear

El índice de compatibilidad léxica nace a partir de los estudios realizados por Ávila Muñoz y Villena Ponsoda (2010, 2011) y Ávila Muñoz y Sánchez Sáez (2011). Los autores parten de la hipótesis de que el índice de disponibilidad léxica permite conocer y acceder al vocabulario del grupo concreto que ha realizado el estudio y, por ende, al prototipo léxico comunitario creado para cada centro de interés. Este hecho se observa al contrastar los índices de disponibilidad de una misma palabra en estudios diferentes ya que se ven alterados, como se expuso en la fórmula López y Strassburger (1987), por el número de informantes, número de menciones y posiciones de la palabra.

¹⁰⁵ <http://gramatica.usc.es/proxectos/dispogal/?lang=es&id=descripcion> [Fecha de última consulta el 9 de abril de 2015].

¹⁰⁶ A través de esta página se puede solicitar el alta de usuario en Dispogal: <http://gramatica.usc.es/dispogal/> [Fecha de última consulta el 9 de abril de 2015].

	Vocablo	Disponibilidad	Frecuencia	Frecuencia acumulada	% Aparición	Rango
Málaga 2010	hígado	0,24165	1,87	36,364	49,534	14
Estudio actual	hígado	0,09071	0,824	73,948	21,637	39

Tabla 2.4. Ejemplo de los resultados del vocablo *hígado* en la investigación de Ávila Muñoz y Villena Ponsoda (2010) y el presente estudio

En la Tabla 2.4, se observa que la palabra *hígado* tiene una disponibilidad de 0,24165 en el estudio realizado en Málaga sobre una población estratificada (Ávila Muñoz y Villena Ponsoda, 2010) mientras que esa misma palabra en el presente estudio tiene una disponibilidad de 0,09071. Se podría decir que ese vocablo es más disponible en el primer estudio porque su accesibilidad es mayor y, por tanto, ha venido a la mente del hablante en más ocasiones y en posiciones más adelantadas. Asimismo, en el primer caso, *hígado* se encuentra más cercana al núcleo del prototipo. Ocupa el rango 14 en el listado de palabras ordenadas según el índice de disponibilidad mientras que, en el segundo caso, ocupa el puesto 39. La diferencia en el porcentaje de aparición de ambos términos es igualmente significativa: en el primer estudio el vocablo ha sido mencionado casi por la mitad de los hablantes mientras que en el segundo no llega a una cuarta parte.

La estructura del léxico en cada centro de interés vendrá determinada por la asociación entre las diferentes unidades léxicas y su grado de accesibilidad. Según Ávila Muñoz y Sánchez Sáez (2011: 257) «es necesario establecer un grado de ‘compatibilidad’ entre, por un lado, cada uno de los vocablos aportados por los hablantes y, por otro, el concepto que para la comunidad estudiada representa cada uno de los centros de interés propuestos». Para establecer diferentes grados de compatibilidad o accesibilidad entre un elemento y su conjunto se basan en el concepto matemático de ‘conjunto difuso’.

El primer paso para calcular el índice de compatibilidad consiste en determinar el espectro de un centro de interés mediante un conjunto difuso a través de un modelo cuantitativo similar al de la ley de Zipf. En este caso, (t) hace referencia a la valoración de cada término obtenida a partir de la división del número que ocupa en la lista un vocablo (n) y una constante para cada caso (k) que representa el tamaño relativo de los conjuntos de términos que superan un nivel de compatibilidad.

$$t = \frac{k}{n}$$

A continuación, construyen el modelo colectivo basado en el valor de nula disponibilidad para todos los términos a partir de la siguiente ley probabilística:

$$a + b - a \cdot b$$

donde

a se refiere a la disponibilidad de un término

b indica la disponibilidad personal de ese término

El índice de compatibilidad muestra la accesibilidad de los vocablos en un rango que oscila entre 1 y 0. En el Gráfico 2.1, las diferentes palabras están representadas en un continuo en el que se destacan los primeros puestos del rango por contar con un índice de compatibilidad cercano a 1. Se constata que a partir de un cierto punto las unidades léxicas pierden compatibilidad y apenas son accesibles a los hablantes.

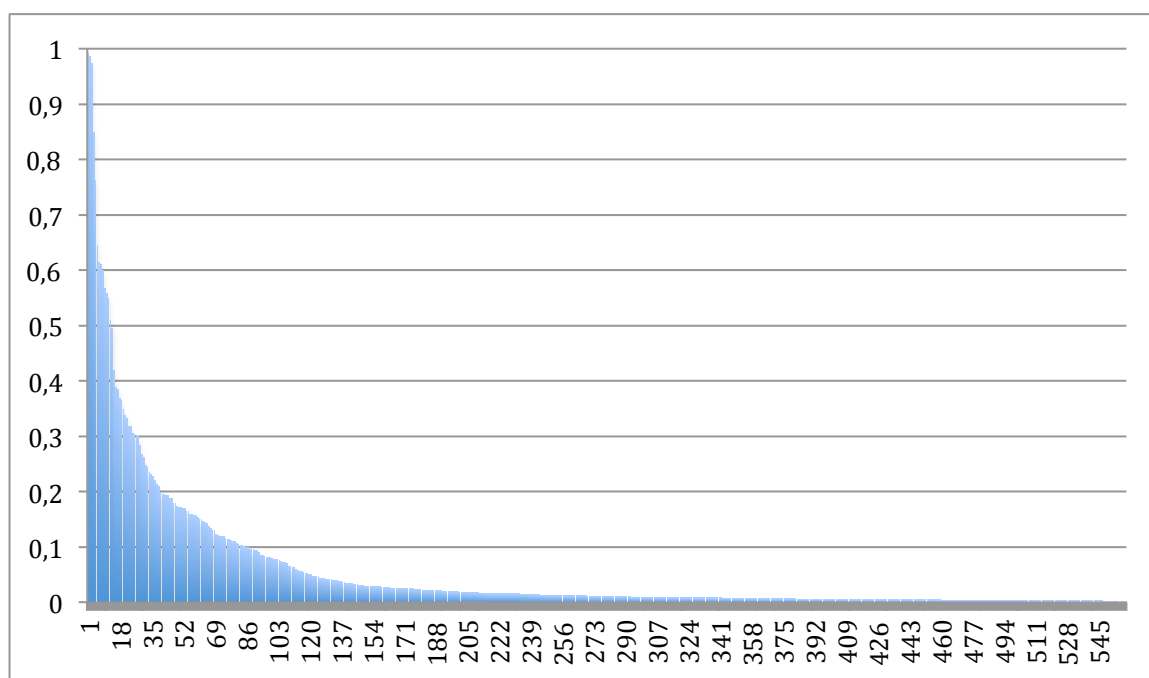


Gráfico 2.2. Número de vocablos ordenados según el índice de compatibilidad en español. Centro de interés 09.
Ordenadores e internet

El índice obtenido sirve de base para dos análisis posteriores que permitirán seleccionar el léxico más compatible y conocer la compatibilidad de cada vocablo de la secuencia léxica aportada por un informante. Por un lado, el planteamiento

expuesto permite aplicar la herramienta FEV (Fuzzy Expected Value) para determinar el léxico relevante. Esa selección del vocabulario vendrá marcada por criterios de corte que establecerán, por niveles, la mayor o menor pertenencia de los vocablos al núcleo, es decir, su nivel de prototipicidad¹⁰⁷. La Tabla 2.1 muestra cómo se ha organizado el vocabulario del continuo anterior en seis niveles según la compatibilidad de los vocablos. En total se han destacado 69 vocablos que suponen un 12,32 % de todos los vocablos aportados por los informantes. Por debajo del índice de compatibilidad de 0,13 se sitúan los vocablos que solo han aportado algunos informantes y, por tanto, que no son representativos del léxico disponible de una comunidad de hablantes.

Nivel de compatibilidad	Proporción	Nº de vocablos	Vocablos
0,85	0,54	3	ratón, pantalla, teclado
0,75	0,89	5	portátil, página web
0,6	1,61	9	ordenador, monitor, Facebook, correo (electrónico)
0,41	2,68	15	torre, impresora, red social, disco duro, programa, archivo
0,24	5,89	33	Tuenti, internet, software, cable, web, hardware, pendrive, Google, blog, chatear, buscador, Windows, escáner, USB, chat, PC, red, wifi
0,13	12,32	69	router, altavoz, módem, CD(-ROM), memoria, descargar, virus, placa base, tecla, descarga, conexión, ordenador de sobremesa, campus virtual, carpeta, navegar, TIC, pc, (Microsoft) Word, sistema operativo, documento, Twitter, procesador de textos, contraseña, DVD, foro, e-mail, videojuego, ordenador personal, información, ventana, CPU, antivirus, batería, netbook, Wikipedia, servidor

Tabla 2.5. Niveles de compatibilidad de los vocablos en español. Centro 09. *Ordenadores e internet*

Por otra parte, este índice de compatibilidad es empleado para conocer cuáles son los procesos de reentrada en el núcleo cuando un hablante realiza la prueba de disponibilidad. El núcleo es considerado el nombre del centro de interés, es decir, la palabra a partir de la cual se activa un vocabulario determinado en la mente del hablante. El orden de producción de las palabras de cada lista puede estar más o menos asociado al vocabulario colectivo. De esta forma, cuando el hablante escribe una palabra con un alto índice de compatibilidad se acerca al núcleo del centro de interés. En cambio, cuando realiza una asociación de tipo personal y, por tanto, no compartida por la mayoría de los hablantes, se aleja de ese núcleo.

¹⁰⁷ Cf. Ávila Muñoz y Sánchez Sáez (2010, 2011a, 2011b, 2014) para una revisión más pormenorizada del tema.

Tras analizar diferentes sujetos, Ávila Muñoz y Sánchez Sáez (2011: 274-277) distinguen cuatro tipos de patrones dominantes de reentrada al núcleo:

- a) Patrón centralizado. Se produce cuando los informantes solo escriben palabras muy compatibles, es decir, muy cercanas al núcleo.

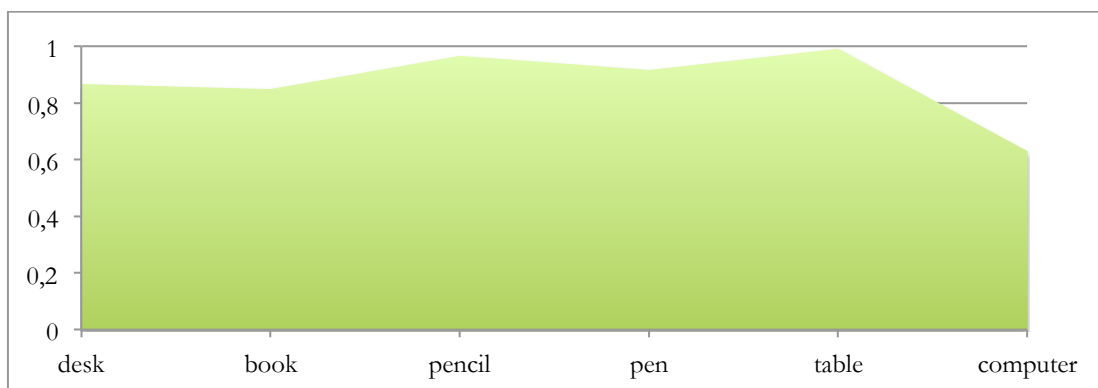


Gráfico 2.3. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 145. Centro de interés 04. *The school: furniture and school material*

- b) Patrón de reentrada. Es el patrón más frecuente observado y se denomina de ‘reentrada’ porque una vez que los informantes han aportado una serie de palabras centralizadas, tienden a alejarse del núcleo. A continuación, vuelven a escribir palabras más centralizadas que lo llevan a sumergirse de nuevo en el núcleo.

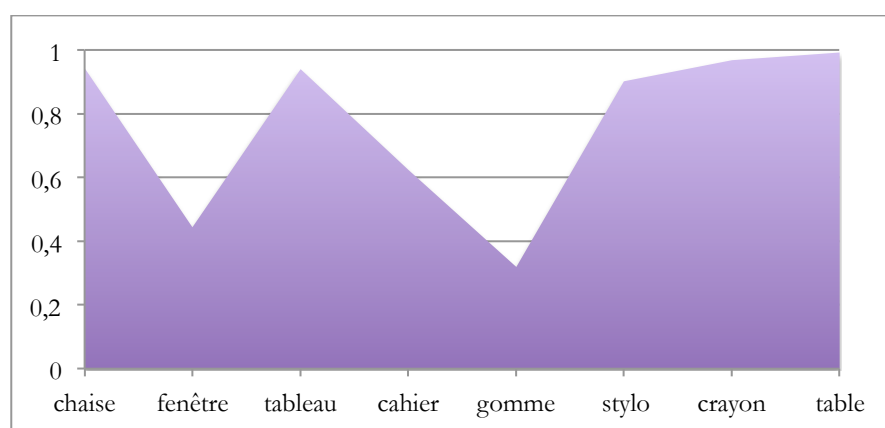


Gráfico 2.4. Muestra de compatibilidad de vocablos. Patrón de reentrada. Informante 170. Centro de interés 04. *L'école: meubles et matériel scolaire*

- c) Patrón de desplazamiento y desviación. Este modelo se observa cuando un hablante, a partir de un determinado momento, comienza a producir palabras que se alejan del núcleo (palabras técnicas o especializadas, asociaciones individuales, etc.).

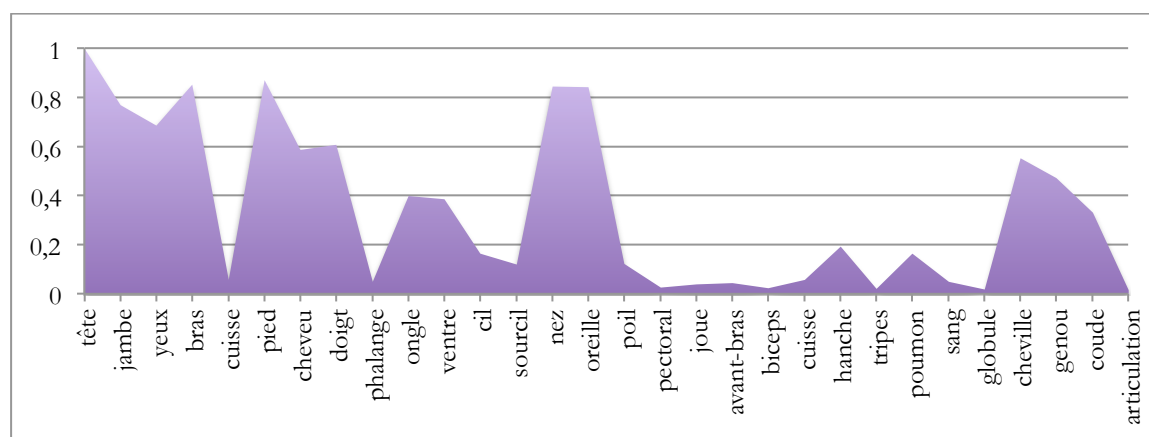


Gráfico 2.5. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 032. Centro de interés 04. *Parties du corps humain*

- d) Patrón de desviación sistemática. Sucede cuando un individuo empieza a escribir palabras menos prototípicas de forma paulatina, desviándose de forma sistemática del núcleo. Este es el patrón menos frecuente porque el hablante, una vez que se ha distanciado del núcleo, tiende a volver al mismo en lugar de seguir alejándose.

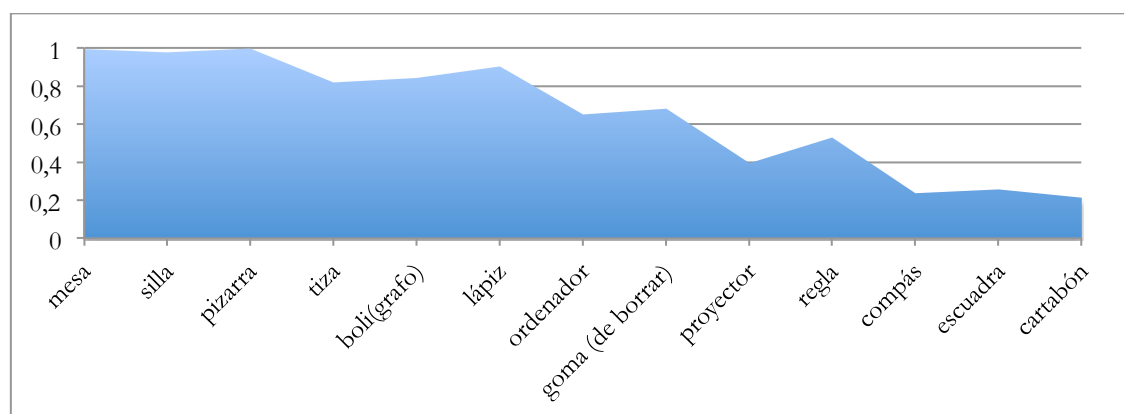


Gráfico 2.6. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 163. Centro de interés 01. *Partes del cuerpo humano*

2.5.3. Índice corregido de amplitud léxica

El índice de amplitud léxica hace referencia al número de palabras aportadas por cada informante y es el resultado de la suma del total de unidades léxicas aportadas en cada centro de interés. En cambio, Ávila Muñoz y Villena Ponsoda (2010: 184) van más allá y distinguen cuatro tipos de palabras según su pertenencia al centro de interés:

- N1: se refiere a asociaciones nucleares, es decir, a palabras que están directamente relacionadas con el centro de interés. Por ejemplo: en el centro de interés *La ropa*, una asociación directa podría ser *pantalón*.
- N5: se trata de vocablos que se alejan del núcleo pero que son compartidos por el colectivo. Por ejemplo: *collar* no designa directamente a una prenda de vestir pero es un complemento que se suele asociar de forma general con la ropa.
- N25: en este tipo se sitúan las palabras derivadas a partir de otras. Por ejemplo: *pantalón de pinzas*, *pantalón chino*, etc. La palabra nuclear sería *pantalón*, pero las compuestas mencionadas se consideran palabras derivadas porque añaden información.
- N0: en esta categoría se localizan las asociaciones arbitrarias que dependen de cada hablante. Por ejemplo: un informante puede llegar a mencionar vocablos como *dieta* o *adelgazar*, que no tienen cabida dentro del centro de interés en cuestión, en este caso, *La ropa*.

A partir de esa clasificación, Ávila Muñoz y Villena Ponsoda (2010) establecen la creación de una fórmula para calcular el índice corregido de amplitud léxica (ICoAL). Este índice pondera de forma diferente los cuatro tipos de palabras para obtener un referente que muestre no solo el número de palabras de un informante sino también la relación con el centro de interés. Así, el valor de una palabra nuclear será 1, de una asociación colectiva 0,5, de una derivación 0,25 y de una asociación individual 0.

$$ICoAL = N1 + (N5 \cdot 0,5) + (N25 \cdot 0,25)$$

En el ejemplo, se observa que el informante 21 ha escrito un total de 294 palabras, de las cuales 268 son nucleares, 19 son asociaciones colectivas, 6 derivaciones y 1 asociación individual. En cambio, el sujeto 34 ha aportado más palabras en total, 298, pero el ICoAL es inferior porque ha mencionado menos palabras nucleares.

$$ICoAL \text{ del informante 21} = 268 + (19 \cdot 0,5) + (6 \cdot 0,25)$$

$$ICoAL \text{ del informante 34} = 260 + (27 \cdot 0,5) + (3 \cdot 0,25)$$

Informante	N0	N25	N5	N1	ICoAL	IAL
21	1	6	19	268	279	294
34	8	3	27	260	274,25	298

Tabla 2.6. Informantes con mayor ICoAL e IAL en lengua española

En la fórmula se pueden identificar tres tipos de palabras ya que no tendría sentido multiplicar las palabras fruto de asociaciones arbitrarias (N0) por 0. A diferencia de los estudios clásicos de disponibilidad léxica, este análisis tiene en cuenta no solo el número total de palabras por informantes sino también la nuclearidad de ellas. Se podría decir que este índice es capaz de evaluar la cantidad y calidad de las palabras aportadas por un hablante.

2.5.4. Índice corregido de descentralización léxica

Ávila Muñoz y Sánchez Sáez (2010) relacionan el índice de compatibilidad con la prototipicidad que presenta el léxico en un determinado centro. En cambio, el índice de descentralización léxica lo vinculan a la ‘especificidad’ del vocablo en contraposición con su centralidad. Para medir la especificidad de un vocablo, basta con restar la unidad al índice de compatibilidad del mismo. Así, si la compatibilidad de una palabra es de 0,9 su especificidad o descentralidad será el restante que le falta para pertenecer al núcleo, es decir 0,10. En la fórmula, m hace referencia a la compatibilidad del centro de interés c y al vocablo en cuestión x :

$$1 - m_c(x)$$

Para calcular el índice de descentralización (IDD), se toma como referencia el cálculo de la compatibilidad de cada término con respecto a su centro de interés. Por ello, al igual que en el caso anterior, se ha empleado el cálculo de la suma probabilística y se han ponderado las descentralizaciones iniciales con un valor k de 0,2.

$$IDD = 1 - \prod_x \left(1 - \left(k \cdot (1 - m_c(x)) \right) \right)$$

Sin embargo, al aplicar esta fórmula todos los vocablos son computados con el mismo valor, 1. Esto quiere decir que una palabra fruto de una asociación individual, como puede ser *trabajo* en el centro de interés *La ropa*, contará con un índice de descentralización bastante elevado. Para subsanar este error, Ávila Muñoz y

Sánchez Sáez (2011) optaron por incluir los valores asignados a cada vocablo según su nuclearidad¹⁰⁸. De esta forma, la palabra *trabajo* en el centro de interés de *La ropa* es considerada una asociación individual ($N = 0$), razón por la cual no computará en el cálculo del índice de descentralización y, por tanto, no será considerada específica para ese centro de interés.

Tras aplicar la revisión comentada, el índice pasa a denominarse índice corregido de descentralización léxica. Este parámetro mide la especificidad de cada vocablo con respecto a su centro de interés. El carácter específico del índice vendrá determinado no solo por el número de veces que aparece, sino también por su posición y relación semántica al núcleo.

En resumen, el IDD es una medida de la capacidad léxica individual obtenida a través de la descentralización del vocabulario actualizado en un centro de interés determinado. La ventaja de este índice es que recoge y valora la compatibilidad o coincidencia de los elementos léxicos de un modo simple y objetivo (Ávila Muñoz y Sánchez Sáez, 2010: 81).

2.5.5. Índice corregido de identificación de tecnicismos

La variable Pident, definida en Ávila Muñoz y Villena Ponsoda (2010: 192) como la «Proporción de Identificación de Tecnicismos», hace referencia a los resultados de la prueba del léxico especializado explicada en el Apartado 3.4.3. Para calcular esta variable, basta con sumar el número de respuestas correctas de cada informante, asignándoles el valor de 1 a cada una. Sin embargo, en este estudio se propone la variable denominada «índice corregido de identificación de tecnicismos» (ICIT) que tiene en cuenta las respuestas correctas, incorrectas y no contestadas. En primer lugar, es necesario conocer la fórmula estadística que permita averiguar qué ponderación tiene que tener cada tipo de respuesta para que un informante que contestase a todas las preguntas al azar obtuviese una puntuación de 0 en el ICIT. Para ello, se han seguido las pautas comunes a cualquier examen tipo test, similar a la prueba de léxico, a través de la fórmula y partes siguientes:

¹⁰⁸ N1 (palabras nucleares) = 1; N5 (asociaciones colectivas) = 0,5; N25 (palabras derivadas) = 0,25; N0 (asociaciones individuales) = 0.

$$\sum_{i=1}^N n_i$$

- Un número concreto de preguntas (N). En este estudio cada palabra especializada se podría considerar una pregunta, siendo un total de 18.
- A cada pregunta se le asigna una puntuación. La (n_i) se refiere a la i -ésima pregunta, es decir, a la puntuación por preguntas (la i significa número de).
- Para cada pregunta existe un número de respuestas posibles (M) pero solo una o varias son correctas.
- La puntuación máxima (P) se consigue acertando todas las preguntas. Normalmente, el sistema de calificación es sobre 10, por lo que P es igual a 10.
- Teniendo en cuenta que 10 es la puntuación (P) y que el cuestionario tiene 18 preguntas (N), cada respuesta correcta tendrá el valor de $\frac{P}{N}$, es decir $\frac{10}{18} = 0,56$.
- Las preguntas no contestadas no se tendrán en cuenta en el cómputo ni positiva ni negativamente.
- ¿Cuál sería el valor que se resta por cada respuesta incorrecta? A continuación, se expone la fórmula aplicada para calcular el valor de X , es decir, de la penalización por cada respuesta errónea.

Teniendo en cuenta que cada palabra puede tener hasta nueve posibles respuestas, la probabilidad de equivocarse es de $\frac{1}{8}$ (si se le resta el caso de la respuesta correcta, el resultado es 8). A continuación, se multiplica el valor de una respuesta correcta, en este estudio $P = 0,56$ por el resultado de $\frac{1}{8} = 0,125$ que da como resultado 0,07. Por tanto, esta última cifra deberá ser sustraída al Pident cuando el informante cometa un error.

Para demostrar los valores expuestos, se presenta la función de la distribución de probabilidad de una variable aleatoria por medio de la cual se define la probabilidad de que un suceso ocurra tomando como referencia el conjunto de todos los sucesos o casos. Dentro de la distribución de probabilidad se distinguen dos tipos:

- Variable continua: como su nombre indica, puede tomar cualquier valor a lo largo de un continuo o intervalo.
- Variable discreta: solo puede tomar valores dentro de un conjunto numerable, es decir, no acepta cualquier valor sino solo aquellos que pertenecen al conjunto. En estas variables se dan de modo inherente separaciones entre valores observables sucesivos.

En este estudio, la función de probabilidad que corresponde es la variable discreta ya que la variable no tiene un número infinito de valores, sino solo aquellos que pertenecen a un conjunto. Por ejemplo, en la variable Pident ningún informante podrá tener un valor mayor de 10 ni valores que muestren un continuo de 9,96, 9,97, 9,98, ya que de forma inherente se producen separaciones entre valores observables sucesivos.

Como se indicó previamente, a cada palabra le corresponde una respuesta correcta, por lo que la posibilidad de acertar es $\frac{1}{M}$ y, consecuentemente la posibilidad de equivocarse será $1 - \frac{1}{M} = \frac{M-1}{M}$. La fórmula sería la siguiente:

$$n_i \begin{cases} \frac{P}{N} \text{ con probabilidad } \frac{1}{M} \\ -X \frac{P}{N} \text{ con probabilidad } \frac{M-1}{M} \end{cases}$$

$$n_i \begin{cases} \frac{10}{18} \text{ con probabilidad } \frac{1}{9} \\ -X \frac{10}{18} \text{ con probabilidad } \frac{9-1}{9} \end{cases}$$

Una vez calculada la distribución de probabilidad de una variable aleatoria discreta, es necesario calcular la esperanza matemática o valor esperado, representado con la letra E , para hacer una mejor predicción. De esta forma, se obtendrá el promedio ponderado de todos los valores posibles. En este estudio, se traduciría en la predicción de la respuesta de un informante a una pregunta cualquiera y, concretamente, en la identificación correcta de una palabra léxica especializada.

$$E[n_i] = \frac{P}{N} \frac{1}{M} + \left(-X \frac{P}{N} \right) \frac{M-1}{M}$$

$$= \frac{P}{N} \left(\frac{1}{M} - \frac{X(M-1)}{M} \right) = \frac{P}{NM} (1 - X(M-1))$$

En el caso de la esperanza matemática para toda la prueba, el resultado sería el siguiente:

$$\begin{aligned} E[NOTA] &= E \left[\sum_{i=1}^N n_i \right] \\ &= NE[n_i] \\ &= \frac{P}{M} (1 - X(M - 1)) \end{aligned}$$

Por último, se calcula el valor de X para saber qué valor hay que restar a cada respuesta incorrecta. La fórmula comienza con el objetivo primero de la creación del ICIT: si un informante responde a todas las preguntas al azar su puntuación será de 0.

$$\begin{aligned} E[NOTA] = 0 \rightarrow 0 &= \frac{P}{M} (1 - X(M - 1)) \rightarrow 1 = X(M - 1) \rightarrow X = \frac{1}{M - 1} \\ E[NOTA] = 0 \rightarrow 0 &= \frac{10}{9} (0,56 - 0,07(9 - 1)) \rightarrow 0,56 = 0,07(9 - 1) \\ &\rightarrow 0,07 = \frac{0,56}{9 - 1} \end{aligned}$$

Una vez realizados los cálculos estadísticos oportunos, se concluye que para obtener unos resultados ponderados de la variable ICIT, es necesario asignar los valores expuestos al principio. Es decir, sumar 0,56 puntos por cada respuesta correcta y restar 0,07 por cada respuesta incorrecta. En la **Tabla 2.7** se han incluido las respuestas de cuatro informantes para comparar los resultados según el Pident y el ICIT. Los informantes 137 y 59 han asociado todas las palabras de forma correcta exceptuando una. En el primer caso, el informante ha dejado la respuesta en blanco y en el segundo la ha contestado de forma errónea. El Pident de ambos sujetos es igual porque no da cuenta de la respuesta incorrecta o no contestada pero el ICIT es diferente: 9,52 frente a 9,45.

En cuanto a los informantes 100 y 169, han asociado correctamente la mitad de las palabras. El informante 100 ha dejado 8 respuestas en blanco y ha realizado una relación incorrecta mientras que el informante 169 ha contestado la otra mitad de forma incorrecta. *A priori*, parece que el informante 100 ha realizado solo la parte de la prueba que conocía mientras que el sujeto 169 ha contestado a todo, con independencia de si había identificado el vocablo o no, razón por la cual es posible

que haya acertado algunas respuestas al azar. Este hecho se refleja en los resultados ya que el ICIT del informante 100 es 4,97 y del informante 169 es de 4,41.

Informante	Respuesta correcta	Respuesta en blanco	Respuesta incorrecta	Pident	ICIT
137	17	1	0	9,44	9,52
59	17	0	1	9,44	9,45
100	9	8	1	5	4,97
169	9	0	9	5	4,41

Tabla 2.7. Ejemplos de respuestas en la prueba del léxico especializado en español y resultados del Pident e ICIT

2.6. Aplicaciones de la disponibilidad léxica

Los estudios en este ámbito surgen de la búsqueda de una aplicación práctica concreta: la selección del léxico para la enseñanza del francés (Gougenheim *et al.*, 1964). Asimismo, los precursores franceses realizan comparaciones y análisis en los que se tienen en cuenta variables sociales, como el sexo o la procedencia, ligados a la incorporación de nuevas disciplinas, como la dialectología o la sociolingüística. Numerosos autores, como López Morales (1995-1996: 249-254; 1999a: 19-26); Carcedo González (1998b: 53-63), Samper Padilla (2003: 213-219), Samper Padilla, Bellón Fernández y Samper Hernández (2003: 105-119) y Ayora Esteban (2007: 2367-2378), destacan el carácter multidisciplinar del léxico disponible y las oportunidades que ofrece. En los párrafos sucesivos se abordan algunas de las principales aplicaciones de la disponibilidad léxica en el ámbito de la dialectología, la etnolingüística, la sociolingüística y la lingüística aplicada a la enseñanza de lenguas.

Desde la perspectiva etnolingüística¹⁰⁹, la disponibilidad puede ayudar a describir las características de una cultura o sociedad y época histórica. Las palabras más disponibles de los hablantes dan cuenta de sus hábitos alimenticios, de la forma de divertirse, de las profesiones y problemas de la zona y demás componentes culturales en el momento preciso de la encuesta. En algunos estudios se han adaptado los centros de interés introduciendo *el mar* (Ávila Muñoz, 2006; Ayora Esteban, 2006; etc.), *el olivo y el aceite* (Ahumada, 2006) o cambiando *calefacción* por *aire acondicionado* como realiza López Morales (1999) en Puerto Rico. Asimismo, se ha criticado que algunos centros de interés tales como, *Trabajos del campo y del jardín*,

¹⁰⁹ Vid. apartado 6.6.2 para una reflexión más profunda.

tendrían que ser actualizados y reemplazados por otros que definan mejor las costumbres actuales y que presenten una mayor producción léxica.

En cuanto a la dialectología, Gougenheim *et al.* (1964) ya encontraron diferencias en los listados del léxico disponible en los listados de los diferentes departamentos franceses analizados. No obstante, Mackey (1971a) es el primer investigador que describe y analiza las semejanzas y diferencias entre dos sintopías: el léxico de Francia y de Acadia (Canadá). A partir de ahí, son considerables los estudios que comparan, tanto desde el punto de vista cualitativo como cuantitativo, los léxicos disponibles de diferentes zonas. En el ámbito hispano, López Chávez (1992) contrasta, por primera vez, el léxico de diferentes lugares: Madrid, Gran Canaria, República Dominicana y Puerto Rico.

En el ámbito de la psicolingüística, destacan algunas investigaciones que tratan de conocer cómo se organizan las palabras en la mente y cómo se activa el vocabulario disponible. Cañizal Arévalo (1991) analiza las relaciones entre las palabras del centro de interés *Partes del cuerpo* y concluye que las palabras se organizan en torno a un núcleo formando redes o constelaciones con un índice de disponibilidad similar. López Chávez (1993) intenta conocer cómo se organiza el lexicón mental a través del análisis de siete muestras de escolares de diferentes edades. Echeverría (1991) se interesa en las relaciones entre las palabras y desarrolla el índice de cohesión y contribuye a la creación del DispoGrafo, una herramienta computacional para contabilizar las relaciones más frecuentes entre las palabras y conocer cómo se organizan (Echeverría *et al.*, 2008). Hernández Muñoz (2006) su describe los fundamentos teóricos que describen la producción del léxico disponible y se centra en los procesos cognitivos.

En esta misma línea, podrían enmarcarse dentro de la psicolingüística no solo las pruebas que siguen la metodología de la disponibilidad léxica del Proyecto Panhispánico sino todas aquellas basadas en la fluencia verbal y que se ponen en práctica en diversos ámbitos. A modo de ejemplo, cabe comentar que en la descripción de afasia en una enciclopedia médico-quirúrgica (Viader, Lambert, de la Sayette, Eustache, Morin, Morin y Lechevalier, 2002), se menciona la aplicación de algunas pruebas típicas como la disponibilidad léxica o la fluencia verbal. En general,

el ejercicio consiste en mencionar durante un tiempo determinado (frecuentemente se elige un minuto y medio) palabras de forma oral relacionadas con un campo léxico (nombres de animales) o nombrar palabras que empiecen por una letra determinada. Los *Animales*, precisamente, forman parte de un centro de interés tradicional y es el ejemplo dado en este artículo de la enciclopedia. Por tanto, este hecho da cuenta de los numerosos estudios que incorporan las pruebas de disponibilidad léxica o fluencia verbal y que pueden encontrarse no solo en las investigaciones sobre la afasia sino también en otros campos, como puede ser el alzheimer (Eustache, 1993: 205-227).

La sociolingüística está presente de una forma u otra en todos los estudios de disponibilidad léxica. En Gougenheim *et al.* (1964) se introdujeron dos variables sociales: el sexo y la procedencia (rural o urbana). A partir de ahí, cada estudio ha incluido unas variables u otras según sus objetivos específicos. Los análisis pasan de cuantificar el número de palabras aportadas por una parte de la muestra, como puede ser el léxico del sociolecto femenino y masculino¹¹⁰, hasta realizar análisis estadísticos más complejos que miden la incidencia directa o indirecta de determinadas variables¹¹¹.

En el campo de la lingüística aplicada a la enseñanza de lenguas se sitúan la mayoría de investigaciones sobre disponibilidad léxica. No obstante, con frecuencia los estudios comentan la importancia de la disponibilidad en la enseñanza de idiomas pero se limitan a aportar los inventarios del léxico disponible a modo de léxico básico y, solo en algunas ocasiones, incluyen criterios de corte para establecer una selección. En cambio, otros estudios cotejan el léxico resultante de las pruebas del léxico disponible con el léxico presente en algunos manuales, como ocurre en Benítez Pérez y Zebrowski, 1993; Benítez Pérez, 1994, 1997; Rueda Bernao, 1994; García Marcos y Mateo García, 2000; entre otros.

A través de un Proyecto Fondecyt¹¹² denominado *Metacognición en lengua materna. Software multimedial para su desarrollo* se crean tres programas informáticos destinados a la enseñanza del español como lengua materna. El objetivo es hacer una

¹¹⁰ La mayoría de estudios suelen incluir esta distinción.

¹¹¹ Entre las investigaciones con mayor profundización estadística destacan los estudios de Ávila Muñoz y Villena Ponsoda (2010, 2011).

¹¹² Fondo Nacional de Desarrollo Científico y Tecnológico de Chile.

reflexión sobre los modelos psicolingüísticos del lexicón mental para crear herramientas de aprendizaje que permitan un aprendizaje significativo. Aunque en el «Taller Internacional de Software Educativo TISE'99» celebrado en Santiago de Chile en diciembre de 1999 se presentara una versión preliminar, en 2001 Echeverría dedica un artículo a exponer las características de uno de los programas, *Vocabulario disponible* y menciona la creación de los otros dos: *Actos de Habla* y *Lectura Crítica* (Echeverría, 2001a). Un año más tarde, Echeverría presenta de nuevo el *Vocabulario disponible* junto con otros programas que comparten características comunes como el hecho de ser aplicaciones de Windows multimediales e interactivas y estar centradas en el objetivo didáctico de la enseñanza del español como lengua materna: *Variación Dialectal del Español*, *Tiempo de Lectura*, *Lector*, *Actos de Habla* y *Antex'* 98.

Vocabulario disponible incluye una breve reflexión Echeverría (2001, 2002) sobre las bases neurales del léxico que finaliza con los principios generales planteados por Caramazza (1999: 470). A continuación, reflexiona sobre la búsqueda del modelo psicolingüístico del lexicón mental y se centra en el modelo conexionista. Esta teoría parte de la idea de que la mente está organizada en forma de redes neurales cuya información (nodos) se encuentra conectada con otros nodos mediante uniones nerviosas. El significado de cada nodo viene determinado por las relaciones con otros nodos o palabras dentro de un complejo sistema.

Una vez esbozados los fundamentos teóricos que intervienen en el diseño del programa, se toma como referencia el léxico disponible de una muestra de más de 2000 estudiantes chilenos de educación media (Valencia Espinoza y Echeverría, 1999) sobre 18 centros de interés¹¹³. Por tanto, la prueba de disponibilidad léxica supone la herramienta de selección del vocabulario en este programa de enseñanza del vocabulario del español de Chile¹¹⁴. La aplicación se divide en tres módulos: *Introducción*, *Diagnóstico* y *Tareas*.

En la *Introducción* se presenta el programa y sus principales características y tareas. Asimismo, se abordan los aspectos teóricos relacionados con la disponibilidad

¹¹³ 1) *Las partes del cuerpo*; 2) *Ropa y calzado*; 3) *La casa*; 4) *Muebles*; 5) *Alimentos*; 6) *La cocina*; 7) *Medios de transporte*; 8) *Trabajo de campo y jardín*; 9) *Plantas y animales*; 10) *Juegos y entretenimientos*; 11) *Profesiones y oficios*; 12) *La política*; 13) *La actividad económica*; 14) *Las artes*; 15) *Ciencia y tecnología*; 16) *Mundo espiritual*; 17) *Procesos mentales*; 18) *Problemas del ambiente*.

¹¹⁴ Se seleccionó el vocabulario que se enseña en el colegio privilegiando, por tanto, la norma culta.

léxica y con la forma de organizar las unidades léxicas. Este apartado es de gran utilidad tanto para el estudiante, para fomentar el aprendizaje autónomo, como para el docente, que puede conocer las orientaciones metodológicas en las que se basa el programa.

En el *Diagnóstico*, el estudiante se somete a una evaluación inicial sobre los centros de interés que desee. Ese examen toma como referencia la prueba de disponibilidad léxica: los informantes escribirán todas las palabras que se le vengan a la mente en los centros marcados en un tiempo determinado. Una vez finalizada la prueba, los datos se presentan en un diagrama que incluye la media de palabras aportadas por los estudiantes chilenos en esos centros de interés. Además, se muestra la diferencia entre la situación de partida del individuo con respecto al promedio de la muestra analizada. Asimismo, se marcan los vocablos que no aparecen en los listados del estudiante y que cuentan con un alto índice de disponibilidad en la muestra del léxico disponible recogida.

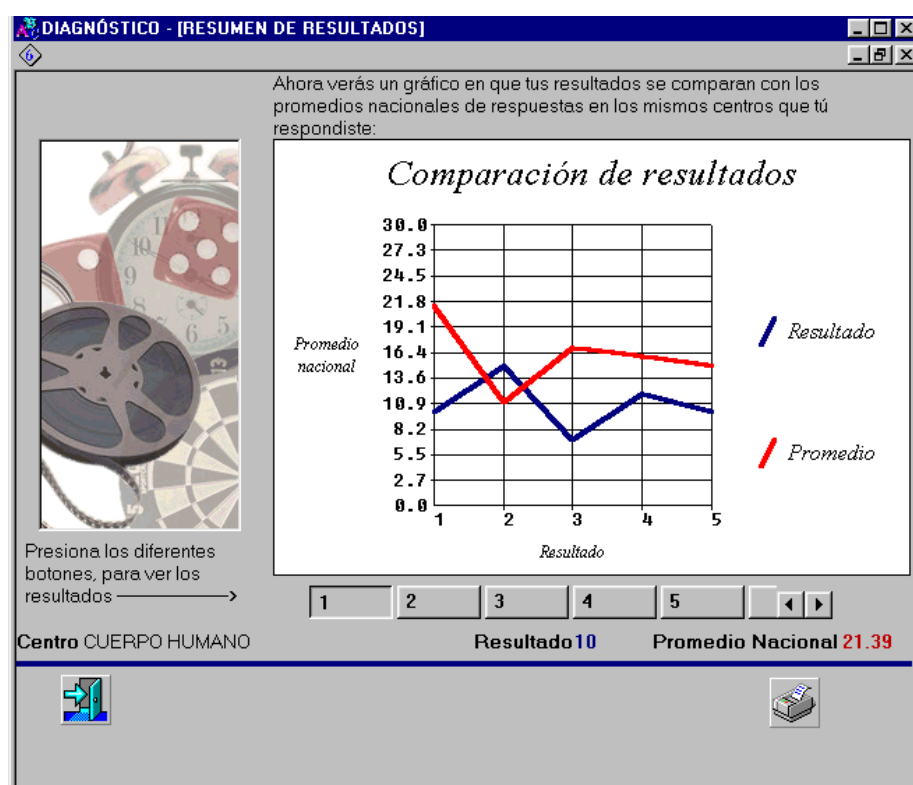


Gráfico 2.7. Estadístico de resultados comparados. Fuente: Echeverría (2001: 94)

En las *Tareas*, el estudiante puede elegir entre nueve tipos de ejercicios diferentes¹¹⁵ y, a continuación, el centro de interés sobre el que quiere realizar la actividad. Además, este programa es editable gracias a la estructura de sistema de autor (*authoring system*) de forma que tanto el docente como el estudiante pueden incorporar nuevos vocablos, actividades, imágenes y vídeos en cualquiera de las nueve tareas.

Aparte de los módulos de aprendizaje, el programa incluye otros dos apartados denominados *Glosario* y *Notas al profesor*. El primero consiste en un buscador de palabras que dirige al diccionario del léxico disponible. El segundo es una interfaz de comunicación que facilita la comunicación entre el alumno y el profesor: permite plantear dudas, comentarios a las actividades realizadas, reforzar algún conocimiento, etc.

Hernández Solís (2006) realiza una propuesta didáctica donde pone en práctica los resultados de las pruebas del léxico disponible realizadas en Zacateca. Entre los objetivos de su tesis doctoral destaca la búsqueda de una aplicación práctica de la disponibilidad léxica (Hernández Solís, 2006: 2). Para ello, realiza una microplanificación la realiza a partir de una fundamentada selección del vocabulario que, en general, proviene de los listados del léxico disponible de los estudiantes universitarios, por considerar que esta muestra de estudio cuenta con el léxico indispensable que deben conocer los estudiantes preuniversitarios. No obstante, al igual que otros autores; reconoce lagunas en el léxico obtenido mediante los centros de interés que completa con vocablos del *Diccionario del Español de México (DEM)*.

La propia autora critica que en otros estudios, como por ejemplo en López Morales (1986), no se realice una propuesta sobre la selección del vocabulario (Hernández Solís, 2006: 6) aunque ella misma admite que presenta un modelo pero no una propuesta acabada (Hernández Solís, 2006: 190). Básicamente, el patrón consta de doce pasos en los que se presentan ejemplos de ejercicios basados en el centro de interés décimo, *La ciudad*. Los pasos son: evaluación de diagnóstico, selección del vocabulario, lectura, ilustración, definición, ilustración, construcción,

¹¹⁵ Asociar una imagen a una de varias palabras, asociar una palabra a una de varias imágenes, mímica, parear un vocablo con su definición, generación de un lexema según sus rasgos semánticos, clasificación de términos, vocablo discordante, exploración en lámina y concentración y memoria.

lectura y corrección, agrupamiento, combinación (las oraciones de base para redactar cada uno de los núcleos), combinación (todas las oraciones de base para redactar el discurso completo) y evaluación final.

El léxico seleccionado corresponde a los 99 vocablos más disponibles de ese centro que se corresponden con el 75 % de la frecuencia acumulada del índice de disponibilidad léxica. Como la mayoría de unidades léxicas son sustantivos, incorpora verbos del *DEM*, que no aparecen en los listados de la disponibilidad léxica de los estudiantes de primaria ni de la universidad, y otro tipo de palabras que incluye en forma de anexo.

2.7. Variables que inciden en el léxico disponible

Desde los inicios de los estudios de disponibilidad léxica los resultados han sido objeto de análisis desde perspectivas diversas, tales como la sociolingüística y la psicolingüística. Los investigadores franceses Gougenheim *et al.* (1964) realizaron una comparación geográfica tomando como referencia cuatro departamentos administrativos del país galo, Dordogne, Eure, Marne y Vendée (Gougenheim *et al.* 1964: 163-164) y una comparación sociológica (Gougenheim *et al.*, 1964: 165-182) basada en dos factores, el medio (rural y urbano) y el sexo.

Las variables sociológicas fueron empleadas por estudios posteriores hasta consolidarse en el programa de referencia en la disponibilidad léxica en la década de los noventa, Lexidisp. Esta aplicación informática permite filtrar vocablos y realizar análisis matemáticos según cinco variables¹¹⁶. Las variables más utilizadas en las investigaciones que se inscriben dentro del Proyecto Panhispánico de Disponibilidad Léxica corresponden al sexo, nivel sociocultural, zona geográfica (urbana o rural) y tipo de centro (público o privado).

En este estudio se analizan las variables sexo, estancias en el extranjero, conocimiento de más de un idioma, hábitos de lectura, ortografía (si lo comparo con el número de palabras), clase social, uso de la lengua en diferentes contextos, exposición a los medios de comunicación, motivación del futuro profesorado y tipo

¹¹⁶ Las variables de referencia son el sexo, el centro escolar (público o privado), la comunidad (urbano o rural), el nivel sociocultural (alto, medio, bajo) y la edad (distinguiendo diferentes generaciones). *Vid.* apartado 3.3.1 para una revisión del funcionamiento del programa.

de estudios universitarios. Ciertas variables han sido analizadas en el ámbito de la disponibilidad léxica mientras que otras suponen una nueva incorporación a estos estudios. Su selección se fundamenta en el ámbito de la lingüística aplicada a la enseñanza de idiomas.

Larsen-Freeman y Long (1991) afirman que existen más de cuarenta teorías diferentes sobre los diferentes factores que inciden en el aprendizaje de una lengua. La mayoría de ellas tienen en cuenta no solo la situación física de aprendizaje, sino también otros factores sociales y reticulares como la edad. Algunos autores como Colletine (2009: 128) se atreven a afirmar que el contexto es una de las variables más importantes¹¹⁷.

El contexto, en sentido amplio, puede abarcar desde la diferente exposición a la lengua (medio endolingüe y exolingüe), hasta las características socioculturales y reticulares del hablante. Existen diferentes clasificaciones al respecto en las que se suelen incluir los siguientes cuatro factores principales que influyen en el proceso de adquisición y aprendizaje de segundas lenguas y lenguas extranjeras: la edad, el método, la aptitud y la actitud (Spolsky, 1969: 272-274). En determinados casos la nomenclatura difiere pero los aspectos considerados suelen ser muy similares. Así, Cenoz Iragui (1993: 29) distingue cuatro tipo de variables clave: individuales, socioestructurales, psicosociales y socieducativas.

La edad es un factor ampliamente estudiado por ser considerado un factor relevante en la adquisición de un idioma (Johnson y Newport, 1989; Long 1990; Newport, 1990; Spadaro, 1998; Scovel, 2000; DeKeyser, 2000; Singleton, 2001; entre otros). Se ha estimado que los niños adquieren con más facilidad un idioma extranjero que los adultos debido a la mayor plasticidad del cerebro.

There is a time in human development when the brain is predisposed for success in language learning. Developmental changes in the brain, affect the nature of language acquisition, and language learning that occurs after the end of the critical period may not be based on the innate biological structures believed to contribute to first language acquisition (Lightbown y Spada, 2006: 68).

¹¹⁷ «One of the most important variables that affects the nature and the extent to which learners acquire an L2 is the context of learning» (Collentine, 2009: 128).

Determinar el momento exacto y el tipo de relación entre adquisición y edad es una cuestión inconclusa y que genera gran controversia (Birdsong, 2005). No obstante, algunos investigadores como Klein (Klein, 1986: 9)¹¹⁸ aseguran que esa facilidad para el aprendizaje de una lengua se extiende hasta la pubertad. La edad adulta conlleva otros beneficios lingüísticos ligados a la abstracción de los términos, la comprensión de estructuras complejas y la memorización de reglas. En general, el adulto adquiere de forma más rápida el vocabulario y es capaz de producir estructuras más complejas que los niños. Según (Lightbown y Spada, 2006: 72), los adultos muestran un desarrollo mayor en las primeras etapas de aprendizaje aunque les resulte muy difícil aprender un idioma con un dominio casi nativo.

El factor educativo, mediante algún método de aprendizaje o clase de aprendizaje, supone en algunos casos la única exposición del hablante a la lengua extranjera. A pesar de que autores como Krashen (*cf.* 1987: 59)¹¹⁹ consideran que no es suficiente para adquirir una lengua, reconocen su importancia como punto de partida para que los estudiantes usen la lengua en el mundo exterior. En este sentido, la mayoría de las investigaciones realizadas en centros educativos con programas bilingües (Gajo y Serra 2002; Braun, 2007) muestran los efectos positivos, no solo en áreas lingüísticas, sino también en el resto de materias. Lyster (Lyster, 2007: 14)¹²⁰ afirma que los programas de inmersión lingüística, con independencia de la modalidad del programa, ayudan a desarrollar la competencia en una segunda lengua.

La aptitud y la actitud hacen referencia a procesos diferentes. Por aptitud se entiende la cualidad o capacidad de un hablante para aprender un idioma. En cambio, la actitud está relacionada con la motivación y su influencia ha sido ampliamente estudiada (Gadner y Lambert, 1972; Gardner, 1982, Castro Calvín, 1991; Dörnyei, 1994a y 1994b). La clasificación tradicional (Baker, 1992, 2001; Finegan, 1994; entre otros) distingue dos tipos de motivación: instrumental, para conseguir un determinado objetivo; e integradora, como medio de formar parte de una sociedad.

¹¹⁸ «In general, it is reasonable to assume that little progress is made after the age of puberty» (Klein, 1986: 9).

¹¹⁹ «The classroom will probably never be able to completely overcome its limitations, not does it have to. Its goal is not to substitute for the outside world, but to bring students to the point where they can begin to use the outside world for further acquisition» (Krashen, 1987: 59).

¹²⁰ «Research has clearly demonstrated that immersion students, regardless of program type, develop much higher levels of second language proficiency than do non-immersion students studying the second language as a regular subject (i. e. for one period per school day)» (Lyster, 2007: 14).

A continuación, se abordan de forma concisa dos factores de aprendizaje que están relacionados con las hipótesis del presente estudio: las estancias en el extranjero y el conocimiento de más de un idioma. Cabe precisar que ambos aspectos han sido poco explorados en el ámbito de la disponibilidad léxica, con lo que los antecedentes bibliográficos son escasos.

2.7.1. Sexo

En la mayoría de los trabajos de disponibilidad léxica las diferencias entre el sociolecto femenino y masculino son muy escasas (Gougenheim *et al.*, 1956; Dimitrijević, 1969; Echeverría *et al.*, 1987; Alba Ovalle, 1995a; Mateo García, 1998; Valencia y Echeverría, 1991; Carcedo González, 2001). En los estudios llevados a cabo por Valencia y Echeverría (1999: 257) y Trigo Ibáñez y González Martínez (2011) se constata que en casi todos los centros de interés el sexo femenino es más productivo que el masculino pero sin llegar a ser una variable significativa.

Gómez Molina (2006) analiza la influencia cualitativa y cuantitativa de esta variable en las encuestas de disponibilidad léxica. El autor parte de la hipótesis de que el diferente comportamiento lingüístico entre hombres y mujeres es de corte sociolingüístico y psicolingüístico. Para ello, analiza la media de palabras por informantes y el índice de cohesión de diferentes estudios¹²¹ realizados a ambos lados del Atlántico. Los centros de interés seleccionados son diecisiete, de los cuales dieciséis corresponden a centros tradicionales y uno nuevo presente en algunas investigaciones¹²² denominado *Los colores*. Gómez Molina (*cf.* 2006: 64)¹²³ concluye

¹²¹ República Dominicana (Alba Ovalle, 1995); Chile (Valencia y Echeverría, 1999); Las Palmas de Gran Canaria (Samper Padilla y Hernández Cabrera, 1997); Salamanca (Galloso Camacho, 2003); Soria (Bartol Hernández, 2004); Castilla la Mancha (Hernández Muñoz, 2005); Cádiz (González Martínez, 2002); Asturias (Carcedo González, 2001); Valencia (en castellano) (Gómez Molina y Gómez Devís, 2004) y Valencia (en catalán) (Llopis Rodrigo, 2005, 2008; Gómez Devís y Llopis Rodrigo, 2010a, 2010b).

¹²² Alba Ovalle, 1995; Valencia Espinoza y Echeverría, 1999; Samper Padilla y Hernández Cabrera, 1997; Galloso Camacho, 2003; Bartol Hernández 2004; Hernández Muñoz, 2005; González Martínez 2002; Carcedo González, 2001; Gómez Molina y Gómez Devís, 2004; Llopis Rodrigo, 2005; Ávila Muñoz y Villena Ponsoda, 2010.

¹²³ «La exposición anterior sobre la incidencia de la variable ‘sexo’ en los estudios de disponibilidad léxica ha demostrado que en aquellas investigaciones donde se han aplicado métodos multivariantes de análisis, el factor sexual actúa con carácter diferenciador en determinados centros de interés, si bien con mayor o menor significación estadística. Dicha covariación con el léxico disponible está motivada, principalmente, por causas de índole sociolingüística (biológicas y patrones socioculturales) y psicolingüística (relaciones asociativas intralingüísticas, preferencias léxicas, etc.)» (Gómez Molina, 2006: 64).

que el conjunto de palabras aportadas por los dos sexos coincide en un alto porcentaje aunque las diferencias entre los sociolectos son relevantes en ciertos centros de interés. Asimismo, aboga por una metodología interdisciplinar desde la cual analizar diferentes aspectos (lexicón mental, memoria semántica y *priming*) presentes en los estudios de disponibilidad léxica¹²⁴.

2.7.2. Estancias en el extranjero

Las estancias en el extranjero son consideradas con frecuencia el mejor método para aprender una lengua, tanto por el profesorado como por el alumnado y, además, existen numerosos estudios que avalan su eficacia¹²⁵. Según Freed (1995a: 4) «[...]study abroad is here considered as a special case of second language acquisition, one which offers a unique opportunity to observe the language learning process from a number of different perspectives». En esa obra, la investigadora recoge una gran variedad de trabajos relacionados con el tema divididos en cuatro bloques temáticos: el primero está dedicado a investigaciones que tienen por objeto predecir y medir los beneficios del lenguaje que reportan los estudios en el extranjero (Brecht, Davidson y Ginsberg, 1995); el segundo comprende cuatro estudios comparativos (Lafford, 1995; Freed, 1995b; Guntermann, 1995 y Huebner, 1995); el tercero recoge tres estudios de corte sociolingüístico (Marriott, 1995; Siegal, 2005 y Regan, 1995); el cuarto consta de tres estudios relacionados con el aprendizaje de la lengua en estancias en el extranjero, es decir, en un contexto de inmersión lingüística (Polanyi, 1995, Miller y Ginsberg, 1995 y Brecht y Robinson, 1995).

En general, la investigadora divide los tipos de estudios relacionados con las estancias en el extranjero en dos grupos: por una parte, investigaciones de corte cuantitativo basadas en resultados provenientes de pruebas que muestran el desarrollo lingüístico; por otra parte, estudios cualitativos que recogen el comportamiento y la percepción de los estudiantes. Sin embargo, Freed (1995: 16)¹²⁶

¹²⁴ Ese análisis multidisciplinar se ha llevado a la práctica en el Capítulo 6 de este estudio.

¹²⁵ Carroll, 1967; Goodwin y Nacht, 1988; Carlson, Burn, Ussem y Yachimowicz, 1990; DeKeyser, 1991; Coleman y Rouxville 1993; Freed, 1995; Pellegrino, 1998; Regan, Howard y Lemée, 2009; Mougeon, Nadasdi, Rehner, 2010.

¹²⁶ «With a few notable exceptions, the finding from the studies to date tell us little about actual language use and serve, therefore, as preliminary explorations of this topic. No study has yet described the range of linguistic

incide en la importancia de seguir investigando en las habilidades que un estudiante adquiere en el extranjero y concluye que ningún estudio ha descrito cómo influyen las estancias en el extranjero en las variables lingüísticas, sociolingüísticas y en los aspectos discursivos.

Ellis (1994: 616)¹²⁷ apunta que el aprendizaje formal de una lengua ayuda a mejorar la competencia en una segunda lengua sobre todo si está asociado a la exposición con el contexto natural. En esta misma línea, Muñoz Lahoz (2002) analiza los factores sociales, pedagógicos y personales que influyen en el aprendizaje de un idioma y afirma que todo el mundo que ha aprendido una primera lengua puede aprender otras lenguas¹²⁸. Por ello, uno de los objetivos del libro consiste en dotar al aprendiz y al enseñante de estrategias de aprendizaje y métodos de enseñanza abordando temas clave con las estancias en el extranjero.

Muñoz Lahoz (2002) considera que a través de las estancias el hablante alcanza mayor fluidez que en un ámbito meramente académico por la exposición al contexto natural. El mayor o menor éxito depende de factores como la edad y la formación. Así, los jóvenes consiguen un aprendizaje lingüístico más rápido que los adultos y niños por contar con mayores recursos lingüísticos y capacidad de reflexión. A su vez, los niños suelen estar más expuestos que los adultos a la lengua y consiguen un aprendizaje más satisfactorio. No obstante, comenta algunos inconvenientes del aprendizaje en un contexto endolingüe, tales como la falta de corrección, que se verían perfeccionados con un aprendizaje formal.

La literatura en este ámbito muestra que la destreza lingüística más estudiada es la expresión oral (Freed, 1995; Freed, Segalowitz, y Dewey, 2004; Lapkin, Hart y Swain, 1995; Lennon, 1990; Juan-Garau y Pérez-Vidal, 2007; Llanes Baró, 2010; Llanes Baró y Muñoz Lahoz, 2009; Yager, 1998). No obstante, existen investigaciones basadas en la fonología (Simões, 1996; Stevens, 2001; Díaz-Campos y Lazar, 2003; Díaz-Campos, 2004; Lord, 2010), el léxico (Milton y Meara, 1995; Ife,

variables (phonologic, syntactic and/or semantic), sociolinguistic and discourse features, that may be influenced as a result of a study abroad experience» (Freed, 1995: 16).

¹²⁷ «In general, then, there is support for the claim that formal instruction helps learners [...] to develop greater L2 proficiency, particularly if it is linked with opportunities for natural exposure» (Ellis, 1994: 616).

¹²⁸ «Todas las personas que han aprendido una primera lengua pueden aprender una segunda y una tercera» (Muñoz Lahoz, 2002: 23).

Vives Boix y Meara, 2000; Dewey, 2008; Foster, 2009; Llanes Baró y Muñoz Lahoz, 2009), la morfología (Howard, 2004, 2006, 2009) y la expresión escrita (Pérez-Vidal y Juan-Garau, 2009; Sasaki, 2009)¹²⁹. Por último, también existen estudios que analizan cambios muy específicos, como el uso de los pronombres en francés *tu* o *vous* (Kinginger y Farrell, 2004) y el uso del pronombre de negación *ne* (Sax, 2003), entre otras.

En el ámbito del léxico, Milton y Meara (1995) evalúan la riqueza léxica que han adquirido 53 estudiantes de intercambio en la British University. Muestran que el vocabulario en inglés que adquieren durante esa estancia es mucho mayor que el que pudiesen aprender en sus países de origen, acentuándose en aquellos estudiantes con un nivel básico de inglés. Davie (1996: 75) apoya esos resultados y destaca que se produce un avance considerable en el vocabulario de los estudiantes que pasan un año de estancia en el extranjero.

Ife, Vives Boix y Meara (2000) analizan el tipo de progreso en el vocabulario realizado en diferentes niveles lingüísticos. La población de estudio estaba formada por 36 estudiantes de una universidad británica que estudiaban español como lengua extranjera. Dos tercios de los estudiantes viajaron a España durante un semestre y un tercio durante dos semestres para realizar un intercambio en alguna de las diez universidades españolas con la que la universidad británica mantenía acuerdos. Se aplicó el conocido como «Tree Word Association Test» o «A3VT» (Vives Boix, 1995) antes y después del viaje. La prueba permite analizar el vocabulario en términos de mejora en la organización léxica más que en el simple número de palabras adquiridas. Los resultados muestran que los grupos con un buen dominio de español realizan mejoras léxicas considerables tanto en el número de palabras como en la organización del lexicón mental, contrariamente a lo apuntado en estudios anteriores.

García Laborda y Bejarano (2008) realizan un estudio cuantitativo-cualitativo sobre la actitud de los estudiantes tras haber participado en un programa lingüístico y un programa de intercambio. La muestra está formada por siete estudiantes de Valdosta State University que realizaron un curso de español durante tres semanas en

¹²⁹ Llanes Baró (2011 : 198-205) realiza un completo esquema de los estudios relacionados con evidencias empíricas que muestran la influencia del contexto en el aprendizaje de una L2.

México y 46 estudiantes extranjeros de la Universidad Autónoma de Valencia que se encontraban concretamente en el campus de Gandía. Entre sus conclusiones, cabe destacar la similitud en la percepción que tienen los propios hablantes del aprendizaje adquirido durante la estancia en el extranjero. Por tanto, señalan la necesidad de incidir más en la mejora de la lengua extranjera en el caso del programa Erasmus, ya que al ser bastante más prolongado en el tiempo debería haber obtenido respuestas más satisfactorias. A nivel de corrección ortográfica, Llanes Baró y Muñoz Lahoz (2009) comprobaron que estudiantes después de una estancia de tres a cuatro semanas en el extranjero consiguieron reducir el número de errores léxicos de forma considerable.

Existe un gran abanico de posibilidades de estancias en otros países que se diferencian, generalmente, por el tiempo, el objetivo de la estancia y el contexto que rodea al individuo que la realiza (edad, integración, antecedentes lingüísticos, etc.). La mayoría de los estudios relacionados con este ámbito recogen datos de estudiantes que han realizado algún curso lingüístico durante un periodo corto o estudiantes de intercambio que continúan sus estudios durante algunos meses en otro país (Ellis, 1994; Milton y Meara, 1995; Freed, 1995; García Laborda y Bejarano, 2008) y todos coinciden en los beneficios que aporta en el aprendizaje de una lengua.

En el presente estudio las estancias se han dividido en tres tipos:

- Erasmus. La beca o programa Erasmus, del acrónimo del nombre oficial que se le da en inglés European Region Action Scheme for the Mobility of University Students, es un programa de movilidad organizado por la Unión Europea. Los estudiantes y profesores universitarios pueden cursar estudios en países adscritos al plan diferente al país de origen del solicitante durante un periodo de tiempo determinado y con el reconocimiento académico en el país de origen. Según datos publicados en la página web¹³⁰ de la Dirección General de Educación y Cultura de la Comisión Europea durante el curso escolar 2010/2011 y por décimo año consecutivo, España fue el país que más estudiantes Erasmus recibió, concretamente 30 580. Además, durante ese

¹³⁰Disponible en: http://ec.europa.eu/education/erasmus/statistics_en.htm [Fecha de última consulta el 5 de marzo de 2015].

curso el número de españoles que obtuvieron la beca Erasmus superó en un 2,7 % a la cifra de extranjeros que llegaron a España con este programa. Andalucía fue la comunidad que más estudiantes envió en términos absolutos (7640), concretamente un 21,1 % del total de España (31 427).

- Otros tipos de estudios becados. En esta variable se pueden encontrar desde becas en niveles de enseñanza obligatoria, como pueden ser los cursos de verano hasta becas de posgrado.
- Estudios en el extranjero sin beca. Los cursos intensivos de verano son los más frecuentes en este apartado. Suelen pertenecer a empresas que ofrecen alojamiento, actividades culturales y clases en un idioma extranjero. Por tanto, el contacto de los estudiantes suele ser con hablantes de procedencia diversa, reduciéndose el intercambio comunicativo con hablantes nativos al profesorado y a los habitantes locales en momentos puntuales.

Santos Díaz (2015a) presenta algunos datos descriptivos sobre el impacto de las estancias en el extranjero en el «XXXII Congreso Internacional de la Asociación Española de Lingüística Aplicada» de la muestra de este estudio. Al igual que en estudios anteriores (Milton y Meara, 1995; Vives Boix, 1995; Davie, 1996; et.), destaca que los informantes que realizaron algún tipo de estancia aportaron en promedio un mayor número de palabras y vocablos que el resto, sobre todo aquellos que disfrutaron de la beca Erasmus.

2.7.3. Conocimiento de idiomas

Los estudios que toman como variable independiente el conocimiento de una segunda lengua en el aprendizaje en una lengua son cuantiosos. La bibliografía no es tan amplia cuando se trata del conocimiento de un tercer idioma por tratarse de una situación relativamente reciente (Maghsoudi, 2008). Sin embargo, las conclusiones de los primeros estudios ayudan a explicar, en gran parte, el comportamiento de los hablantes plurilingües.

Muestra de ello es la Hipótesis de la Interdependencia Lingüística propuesta por Cummins (1979). Señala que un estudiante que ha desarrollado previamente habilidades lingüísticas relacionadas con el uso del lenguaje, si tiene la posibilidad de

transferirlas a la nueva lengua, lo hará de manera más o menos rápida. Mientras que un estudiante que no ha desarrollado dichas habilidades con anterioridad tendrá dificultades en el aprendizaje de una segunda lengua. Por ello, es más conveniente que las habilidades lingüísticas se desarrollen primero en la lengua materna para poder tener mejor disposición para el aprendizaje de la segunda lengua. Además, Cummins recoge información de investigaciones realizadas en diferentes partes del mundo donde se ratifica su hipótesis (Cummins, 1991: 70-89): «When children begin the acquisition of a second language (L2), whether in the home or at school, their cognitive resources clearly play a central role in the rapidity and ultimate success with which that language is acquired».

En este sentido, gran parte de los estudios neurolingüísticos y psicolingüísticos que dan prueba de los beneficios cognitivos del bilingüismo, extienden esas ventajas al aprendizaje de una tercera lengua (Jacobsen e Imhoof, 1974; Albert, Obler y Obler, 1978). En 1961, Lerea y Kohut realizan un estudio con 30 hablantes monolingües y 30 bilingües en el que durante 70 minutos realizan cuatro pruebas¹³¹: dos lingüísticas, una de inteligencia y una de personalidad. Los resultados muestran que los estudiantes bilingües aprendieron más rápidamente a asociar palabras del test lingüístico que los monolingües hecho que, a su vez, está correlacionado con una mayor inteligencia.

En cuanto a los estudios con una tercera lengua propiamente dichos, cabe destacar el artículo de Mägiste (1984) donde extrae conclusiones de tres investigaciones que evalúan la adquisición de una tercera lengua en estudiantes inmigrantes. La primera, iniciada por el National Swedish Board of Education, compara los resultados en inglés de estudiantes suecos monolingües y de estudiantes que usan el sueco en casa pero tienen un conocimiento pasivo de su lengua materna (no hay homogeneidad en esta variable ya que depende del país de origen del estudiante). Tras pasar una serie de pruebas estandarizadas en inglés divididas en cuatro partes, se constata que los estudiantes bilingües con un conocimiento pasivo de su lengua materna presentan mejores resultados en inglés que los estudiantes

¹³¹Cuestionario bilingüe, Columbia Mental Maturity Scale, Rogers' Test of Personality Adjustment y Micro Utterance-Association Test (MU-A).

monolingües. En cambio, los resultados de los hablantes que usan diariamente su lengua materna, en lugar del sueco, son ligeramente más bajos.

Swain, Lapkin, Rowen y Hart (1990) demuestran que es errónea la tendencia popular a pensar que los niños no deberían recibir instrucción en una L2 o L3 hasta que no muestren un elevado nivel de conocimientos en la lengua mayoritaria. Aseguran que muchos padres inmigrantes en Canadá son reacios a enseñar la lengua materna a sus hijos porque creen que su nivel en inglés se verá afectado. En cambio, apoyándose en la hipótesis de la interdependencia de Cummins y en sus propios resultados, explican que la tercera lengua favorece a los estudiantes de un programa bilingüe inglés y francés en Toronto. Ese mismo año, Nayak, Hansen, Krueger y McLaughlin (1990) señalan que los hablantes multilingües muestran un mejor rendimiento a la hora de adquirir ciertos contenidos gramaticales que los monolingües.

En 1992, Thomas compara el rendimiento de hablantes bilingües (inglés y español) y monolingües (inglés) que estudian una segunda o tercera lengua. Además, clasifica a los hablantes bilingües en dos grupos diferenciando a los estudiantes que han recibido una instrucción formal en una lengua semánticamente relacionada con la lengua meta de aquellos que lo adquirieron en otro contexto. En general, los resultados muestran que la conciencia metalingüística de los estudiantes bilingües puede aumentar si aprenden una tercera lengua. Así, cuando los anglófonos con conocimiento de español realizan actividades para aprender francés, suelen tener más facilidades que los monolingües. En cuanto a los hablantes bilingües que han recibido clases formales en español, se demuestra que han desarrollado una conciencia lingüística que les ayuda a aprender francés más fácilmente que a los bilingües que han aprendido español en un medio informal.

Keshavarz y Astaneh (2004) realizan una prueba de vocabulario en inglés a una población de estudio formada por hablantes monolingües persas y hablantes bilingües de turco y persa o armenio y persa. Las pruebas muestran que los hablantes nativos de turco y armenio que hablan persa como segunda lengua obtuvieron mejor puntuación en inglés que los hablantes monolingües.

El armenio, el persa y el inglés fueron también las lenguas objeto de estudio de Kassaian y Esmae'li en 2011. Concretamente, realizan un análisis del rendimiento en una prueba de vocabulario y comprensión lectora de 30 hablantes bilingües armenios y persas y 30 monolingües persas. Los informantes, niñas iraníes de 17 a 18 años, cuentan con un nivel intermedio de inglés. Cabe mencionar que Irán posee un sistema educativo centralizado por lo que las pautas metodológicas y los manuales utilizados para el aprendizaje de inglés como lengua extranjera fueron los mismos. Los resultados estadísticos muestran que el bilingüismo está altamente correlacionado de manera positiva con el dominio de vocabulario y la destreza lectora.

Aunque estas investigaciones comparten características con estudios centrados en una segunda lengua, autores como Cenoz (2000: 39-53) considera que se debe hacer una clara distinción entre la adquisición de una segunda lengua y la adquisición multilingüe. Según Clyne (1997: 113) el motivo radicaría en que los procesos de adquisición de una tercera lengua son mucho más complejos, en parte debido a los diferentes contextos de aprendizaje no solo del individuo sino del individuo con cada lengua. De hecho, investigaciones como la de Hsu (2013) muestran diferentes resultados no solo entre hablantes monolingües y bilingües, sino también entre hablantes bilingües y trilingües.

En el ámbito de la disponibilidad léxica, los estudios que incluyen la variable referida al conocimiento de otras lenguas lo hacen para valorar su incidencia en el léxico disponible en español como lengua extranjera. Carcedo González (2000c) analiza la influencia del conocimiento de otras lenguas románicas de los estudiantes finlandeses de español como lengua extranjera. Prueba que aquellos informantes hablan otras lenguas romances presentan una productividad léxica mayor que el resto.

Samper Hernández (2002) enriquece el análisis de esta variable en estudiantes de español como lengua extranjera que realizan una estancia en España. No solo pregunta al informante si conoce otra lengua, sino también con qué nivel (inicial, intermedio, avanzado y superior) y cuál es su lengua materna (inglés, italiano, japonés y otros). Destaca que el factor más influyente en el caudal léxico es el nivel alcanzado en lengua española. Sin embargo, la variable referida al conocimiento de otras lenguas

no presenta resultados significativos y los estudiantes japoneses son los que escriben más palabras por encima de los italianos, ingleses y el grupo denominado «otros». No obstante, la autora advierte que se trata de un estudio exploratorio que cuenta con una muestra reducida y cuyos datos pueden servir de base para investigaciones futuras.

Hugo (2003) también realiza una investigación exploratoria donde analiza una muestra de 84 estudiantes alemanes de la región de Sarre que estudian español en el instituto. El estudio concluye que la variable con mayor incidencia en la productividad léxica es el nivel alcanzado en lengua española, con independencia del número de lenguas que conozca el informante. Sánchez-Saus Laserna (2008) analiza varias variables relacionadas con el aprendizaje de una lengua sobre una muestra de 81 estudiantes extranjeros de español de la Universidad de Cádiz: nivel de español (A, B y C, como reducción de los seis del MCER), lengua materna (inglés, francés, alemán, italiano, finés y polaco), conocimiento de otras lenguas (románica o no románica) y uso del español (alto o bajo). Entre las conclusiones, cabe destacar que los informantes que conocen otra lengua románica tienen mayor producción léxica que el resto.

En su tesis doctoral, Sánchez-Saus Laserna (2011) estudia el léxico disponible de 322 estudiantes extranjeros de español en las ocho universidades andaluzas. Las variables seleccionadas son similares: el sexo, el nivel de español (A, B y C, a partir de los niveles del MCER), las lenguas maternas (alemán, finés, francés, inglés, italiano y polaco), la frecuencia con que hacen uso del español (alta, media y baja) y el conocimiento que tienen de otras lenguas extranjeras románicas (solo en el caso de informantes que no tengan ya una lengua románica como lengua materna y que contempla tres posibilidades: que no conozcan ninguna otra lengua extranjera; que conozcan otra u otras, pero ninguna sea románica, y que conozcan otras lenguas románicas).

Los resultados muestran que aquellos estudiantes que conocen otra lengua extranjera producen la misma cantidad de palabras que aquellos estudiantes que hablan otra lengua románica como lengua materna o incluso producen más palabras. Sánchez-Saus Laserna (2011) matiza los resultados indicando que esa variable está

relacionada directamente con la lengua materna y con el nivel de español y, como no existe el mismo número de informantes en cada variable, los datos pueden verse influenciados.

Frey Pereyra (2008) analiza variables similares en una muestra de 31 estudiantes de español como lengua extranjera: lengua materna (sueco, portugués, noruego, mandarín, japonés, inglés, francés, danés, coreano y alemán), sexo, edad, nacionalidad, nivel de estudios de español, conocimiento de otras lenguas extranjeras y tiempo de residencia en España. En el análisis cualitativo y cuantitativo, el conocimiento de otras lenguas juega un papel secundario ya que la variable con mayor incidencia es el tiempo de residencia en España.

En un trabajo preliminar, López Rivero (2008) analiza el léxico disponible de 43 estudiantes adultos de español como lengua extranjera. Las variables objeto de estudio son el sexo, edad, nivel sociocultural, conocimiento de otras lenguas, lengua materna, años de estudio del español y método de enseñanza del español seguido. La incidencia del conocimiento de otras lenguas sobre el vocabulario disponible se hace patente en el número de palabras aportadas por cada informante, de forma que cuantas más lenguas conoce un informante mayor será su caudal léxico. Este factor se acentúa en los hablantes que conocen tres o más lenguas.

Magnúsdóttir (2012) realiza un estudio con una muestra de 288 escolares de Islandia que estudian español como tercera lengua extranjera en Secundaria. Las variables analizadas son el sexo, la ubicación del centro, tipo de centro, modalidad de estudio y conocimiento de otras lenguas. En cambio, para poder realizar los análisis con Displex, Magnúsdóttir (2012: 71) reduce y simplifica las variables lo que conlleva que no se analice el conocimiento de otras lenguas extranjeras.

García García (2014) estudia las variables sexo, edad, lengua materna, nivel de español, motivación para aprender español, conocimiento de otras lenguas, nivel sociocultural y tipo de estudiante sobre una muestra de 300 estudiantes de español en la comunidad de Madrid. El investigador concluye que la variable relativa al conocimiento de otras lenguas no incide en el caudal léxico de los estudiantes ya que se el promedio de vocablos aportados por los informantes que no conocen otra

lengua extranjera es superior a aquellos que conocen otra lengua y el promedio de palabras es similar en ambos grupos.

En el «XI Congreso Internacional de Lingüística General», Santos Díaz (2014a) presenta algunos análisis descriptivos de la relación de la variable del conocimiento de otras lenguas con el número de palabras y vocablos sobre la muestra del presente estudio. En general, el promedio de palabras y vocablos aportados por los informantes aumenta a la par del número de idiomas que conoce un hablante. No obstante, aquellos informantes que tienen un nivel avanzado en dos idiomas presentan un mayor caudal léxico que los sujetos que conocen dos lenguas con un nivel inicial o medio. Por tanto, de los resultados se desprende que la productividad léxica se ve influida no solo por el número de lenguas que conoce un informante sino también por el nivel de competencia lingüística alcanzado en las lenguas que conoce.

2.7.4. Corrección ortográfica

Las directrices dadas en las pruebas de disponibilidad léxica muestran el carácter secundario otorgado a la corrección ortográfica. Como norma general, en los estudios herederos del Panhispánico, al informante se le pide que escriba todas las palabras que conozca relacionadas con un centro de interés dado. Por ello, a menudo los hablantes escriben todo lo que se les viene a la mente descuidando la forma e incluso el contenido, ya que el objetivo es conseguir la máxima producción léxica.

Según Paredes García (1999: 76)¹³² el hecho de que los informantes no puedan reflexionar sobre la forma correcta de cada palabra, como podría hacerse si no existiese la limitación temporal o esta fuese superior, ayuda a conocer la forma original de la palabra en la mente del individuo. Precisamente por esta razón, es tan relevante desde el punto de vista de la didáctica de la ortografía puesto que permite detectar los errores más frecuentes en una población de estudio y diseñar materiales que hagan hincapié en determinadas reglas ortográficas, excepciones a la norma o

¹³² «Podemos decir que lo que las encuestas nos ofrecen es el estado primario de la situación ortográfica de cada sujeto, la forma básica con la que la palabra aparece representada en la mente del individuo. Esto parece constatarse en el hecho de que proporcionalmente son muy escasas las autocorrecciones o las grafías híbridas (tipo b/vaca), índices de dudas sobre el uso correcto del idioma» (Paredes García, 1999: 76).

palabras específicas. A continuación, se presentan estudios de disponibilidad léxica que han analizado en detalle los errores ortográficos presentes en los listados y que, en general, coinciden en la predominancia de los errores de acentuación sobre el resto.

Paredes García (1999) analiza los errores ortográficos presentes en una prueba de disponibilidad léxica realizada a 484 informantes de segundo ciclo de Educación Secundaria Obligatoria y Bachillerato de Alcalá de Henares. Dadas las características de la prueba de disponibilidad léxica, donde los informantes aportan unidades léxicas simples o complejas, Paredes García (1995: 76) afirma que tan solo es posible conocer los errores más frecuentes en el plano fonético y léxico. En su estudio (Paredes García, 1999: 81), los errores de acentuación ocupan el primer lugar (10 625 palabras) frente a los errores de letra (4619 palabras), los errores léxicos (3477 palabras) y los errores de carácter morfosintáctico (320).

Ávila Muñoz (2007) realiza un estudio exploratorio con estudiantes de 2.º de Traducción de la Universidad de Málaga. Destaca la proliferación de las faltas de ortografía en los estudiantes universitarios puesto que de los 50 informantes encuestados, tan solo 3 no cometieron ninguna falta de ortografía, 14 escribieron menos de 10 palabras incorrectas, 19 entre 10 y 20 faltas y 14 cometieron más de 20 faltas. Además, clasifica las diferentes disortografías en errores de acentuación y ortografía general pero no los cuantifica.

A continuación, analiza la incidencia de determinadas variables (sexo, número de horas semanales navegando por Internet, clase social, libros que el estudiante lee al año, entre otras) con objeto de conocer cuáles son los principales condicionantes de las faltas de ortografía. Ávila Muñoz (2007: 45) concluye que las variables que guardan una mayor correlación son la clase social y los hábitos de lectura. Muestra que cuanto mayor es la clase social del informante y más libros lee al año, menos faltas de ortografía comete en la prueba de disponibilidad léxica. En cuanto al resto de variables, muestra que las mujeres tienden a cometer más faltas de ortografía que los hombres (sin tener en cuenta los errores de acentuación) y que las horas de navegación en Internet no tienen una incidencia negativa en la corrección ortográfica.

El estudio de Saura Rami (2008) selecciona informantes de la muestra aragonesa realizada durante los cursos 1998/99 y 1999/00 formada por 417 estudiantes de 2.º de Bachillerato o COU. Aunque los estudios se realizan en 28 centros de enseñanza, el análisis sobre la ortografía de las encuestas recoge tan solo los datos de 45 alumnos de los IES Lucas Mallada (Huesca), Luis Buñuel (Zaragoza) y Francés de Aranda (Teruel). De las 1394 disortografías detectadas, 940 corresponden a errores de acentuación y 454 a errores gráficos, es decir, que los informantes comenten el doble de faltas de ortografías debido a una inadecuada acentuación

Blanco Domínguez (2011) analiza los errores ortográficos de aquellas palabras con un índice de frecuencia acumulado del 80 % de los estudiantes preuniversitarios gallegos de la obra colectiva editada por López Meirama (2008). Del total de 11 152 errores ortográficos, 5756 corresponden a una inadecuada acentuación, lo que equivale al 51,57 %. Se trata hasta el momento del estudio que analiza de forma más pormenorizada las formas erróneas, distinguiendo entre errores ortográficos, errores fonéticos y lapsus. Por ejemplo, cuando el error se debe a la ausencia de tilde distingue entre palabras: agudas (29,22 %), llanas (15,28 %), esdrújulas (37,80 %), con diptongo (6,02 %), con hiato (11,30 %) y monosílabas (0,38 %). En cuanto a la colocación de la tilde en la sílaba incorrecta, un 68,30 % se debe a palabras agudas, llanas y esdrújulas y, el resto (31,70 %), se localiza en palabras con diptongo o hiato. Para completar el análisis cuantitativo, Blanco Domínguez (2011) añade las palabras que presentan con más frecuencia disortografías y, en determinados casos, su forma incorrecta (*autóbus*).

2.7.5. Clase social

El estrato sociocultural forma parte de la mayoría de las investigaciones de disponibilidad léxica. Los precursores de esta disciplina, Gougenheim *et al.* (1964), incluyeron esta variable en las encuestas pero no la tuvieron en cuenta a la hora de realizar análisis. López Morales (1979) toma como referencia los niveles socioeconómicos de la teoría del déficit de Bernstein (1961) para analizar cualitativa y cuantitativamente su incidencia en el léxico disponible. A pesar de que revela

diferencias en el número y el tipo de palabras aportadas por los diferentes estratos, López Morales (1979: 181) considera que esta teoría no es suficiente para afirmar que los estratos inferiores tengan una menor capacidad cognitiva. Años más tarde, López Morales (1996) refuta la hipótesis de que el diferente caudal léxico de los estratos sociales implique dificultades cognitivas en el estrato con menor léxico.

En general, en los estudios de léxico disponible el nivel sociocultural está conformado por los datos relativos al nivel de estudios y la ocupación profesional de los padres del informante (Echeverría, 1991; Alba Ovalle, 1995a; Samper y Hernández, 1997; Mateo, 1998; Echeverría y Valencia, 1999a; Carcedo 2001; Gómez Devís, 2004; entre otros). En cambio, al igual que en este estudio, algunos autores han incluido los ingresos familiares para mejorar la descripción de esta variable (López Morales, 1999a; Carcedo, 2001; Ávila Muñoz y Villena Ponsoda, 2010, 2011).

Ávila Muñoz y Villena Ponsoda (2010) realizan un completo análisis multifactorial donde evalúan la incidencia de la posición social sobre el léxico disponible y la capacidad de identificación de tecnicismos (Pident). Destacan la relación entre la posición social y el léxico especializado de forma que aquellos informantes provenientes de clase alta son capaces de reconocer y asociar correctamente una mayor cantidad de palabras técnicas con los centros de interés. No obstante, concluyen que la productividad léxica está más asociada a la exposición con los medios de comunicación y al conocimiento del vocabulario especializado. Por tanto, la clase social influye de manera indirecta, a través de la variable Pident, en el léxico disponible.

2.7.6. Hábitos de lectura y exposición a los medios de comunicación

A través de la lectura, el estudiante es capaz de aprender numerosas unidades léxicas de manera incidental (Elley y Mangubhai, 1983; Coady, 2000). En este sentido, existen numerosos estudios prueban el papel fundamental que juega la lectura en el proceso de adquisición y aprendizaje de una segunda lengua y una lengua extranjera (Zahar y Cobb y Spada, 2001; Waring y Takaki, 2003; Horst, 2005; Pigada y Schmitt, 2006; Brown, Waring y Donkaewbua, 2008; entre otros). No obstante, el efecto de

los hábitos de lectura y la exposición a los medios de comunicación no son variables comunes en los estudios de disponibilidad léxica.

El primer estudio de léxico disponible que incorpora una variable relativa a la lectura es llevado a cabo por Mackey (1971a) quien se centra en analizar el tipo de lecturas que realizan los informantes. En el año 2007, Ávila Muñoz relaciona el número de libros que los informantes leen al año con los errores ortográficos que comenten en lengua materna. Demuestra que se produce una correlación significativa entre ambas variables de forma que cuanto más errores comete un informante menos libros lee al año. Ávila Muñoz y Villena Ponsoda (2010) incorporan una variable que recoge el número de libros que el informante lee al año y que divide los resultados en tres grupos: entre 0 y 10, entre 10 y 20 y más de 20. En esta ocasión, la incidencia de esta variable no resulta significativa en el caudal léxico en la primera lengua de los informantes.

En cambio, incorporan una variable que hace referencia a la exposición a los medios de comunicación, dividida según el tipo de programas y el número de horas de exposición, que revela diferencias significativas entre los diferentes grupos. Ávila Muñoz y Villena Ponsoda (2010: 264) afirman que el efecto de la exposición a programas televisivos de consumo (noticias del corazón, telerrealidad, etc.) es curvilíneo: quienes siguen estos programas una vez a la semana presentan puntuaciones más bajas en el continuo de desplazamiento (el cual muestra los hablantes con mayor o menor vocabulario descentralizado) que aquellos con nunca ven este tipo de programas o lo hacen asiduamente. Esa exposición a los medios de comunicación no resulta significativa en el número de faltas de ortografía (Ávila Muñoz, 2007).

2.7.7. Contacto y uso de la lengua extranjera en diferentes contextos

Esta variable se refiere al contacto y al uso de la lengua extranjera por parte del informante: para comunicarse con amigos y familiares, para escuchar música, ver programas de televisión, entre otros aspectos. Es de gran relevancia desde el punto de vista de la lingüística aplicada a la enseñanza de una lengua, ya que es necesario marcar la diferencia entre una situación de enseñanza-aprendizaje endolingüe y

exolingüe para conocer el valor del contexto social, psicológico y pedagógico de cada lengua. El medio endolingüe hace referencia al contexto natural donde la lengua meta se utiliza como lengua vehicular de la zona. En cambio, en el contexto exolingüe el aprendizaje se produce exclusivamente en un contexto pedagógico y, por tanto, alejado del contexto natural (García Santa Cecilia, 1995: 110, 2000: 80; Puren, 1998: 363). En el caso de que un español se trasladase a Francia, aprendería el idioma francés en un contexto endolingüe mientras que si ese mismo sujeto aprende francés en España lo hará en un medio exolingüe, normalmente recibiendo clases formales de ese idioma.

Existen otras distinciones y nomenclaturas para designar el contexto de aprendizaje. Así, Pendanx (1998: 38) distingue entre aprendizaje en una situación formal e informal. La primera hace referencia al aprendizaje institucional que recibe un alumno en un aula, independientemente de si lo hace en su país de origen o en el país de la lengua meta. En cambio, el aprendizaje informal se realiza en un determinado medio social o en el país extranjero.

La variedad de situaciones de aprendizaje pueden llevar a equívoco al suponer que lo que determina el proceso de aprendizaje y adquisición es el lugar en vez del modo. Por ejemplo, si un niño nace en España en el seno de una familia rusa, seguramente adquiera como lengua materna el ruso de forma natural y espontánea aunque no sea la lengua vehicular del país. Por tanto, cada proceso está supeditado a las circunstancias concretas del hablante y de la forma en la que se realice la interiorización de la lengua.

Los vocablos *aprendizaje* y *adquisición* se utilizan a menudo como sinónimos pero según el grado de especialización del mensaje se pueden contraponer. Ello se debe a que en el lenguaje familiar sendos términos designan los procesos de adquisición y aprendizaje de una lengua en tanto que se refieren al proceso de interiorización de una nueva lengua. Sin embargo, en un sentido más purista, el concepto de *aprendizaje* se refiere al contexto de una lengua extranjera y, por ende, al aprendizaje formal: en un aula, normalmente con un profesor, libro, etc. Por su parte,

el término *adquisición* se relaciona con el aprendizaje de una lengua materna¹³³ y de una segunda lengua, ya sea en contexto natural o mixto, es decir cuando además de aprender la lengua por simple contacto, el hablante recibe clases.

Las aclaraciones a los términos presentados llevan consigo la demarcación de *segunda lengua* frente a *lengua extranjera*, términos a su vez frecuentemente confundidos pero con un rasgo diferenciador en el ámbito de enseñanza de lenguas. El aprendizaje de una lengua extranjera (LE) se produce en un contexto donde la lengua no es la nativa de la sociedad, como el caso de un estudiante de árabe en Italia, con lo que suele estar ligado al aprendizaje en el aula. No obstante, la adquisición de la segunda lengua se produce en un contexto en el que el aprendiz está inmerso en la lengua nativa de la mayoría de la población. Este sería el caso del francés para un emigrante chino que se instale en el país galo y adquiera la lengua de forma natural o directa. Cabe destacar que algunos investigadores no respetan esta convención y se refieren a la lengua extranjera de forma indiscriminada para hablar de segunda lengua y lengua extranjera, si bien no extienden el término segunda lengua a ambos casos.

Tal distinción toma como criterio marcador el entorno sociolingüístico en que se realiza el aprendizaje de la lengua meta. Una de las distinciones más extendidas al respecto es la planteada por Stephen Krashen en su conocida «Input Hypothesis and second-language acquisition theory». La primera de las cinco hipótesis expuesta se titula «The Acquisition-Learning Hypothesis» y basa la diferencia entre adquisición y aprendizaje entre lo inconsciente y lo consciente, respectivamente.

There are two ways of developing ability in second languages. 'Adquisition' is a subconscious process identical in all important ways to the process children utilize in acquiring their first language, while 'learning' is a conscious process that results in 'knowing about' the language (Krashen, 1985: 1).

Algunos estudios de disponibilidad léxica en lengua extranjera incorporan variables relacionadas con el uso de la lengua. Escoriza Morera (2003) evalúa el uso del español y del inglés que hacen los gibraltareños preuniversitarios a través de variables como la lengua materna, la lengua utilizada en casa y la lengua utilizada

¹³³ La lengua materna o L1 se entiende como aquella que el niño adquiere en condiciones normales en el seno de su familia, de ahí que se llame también primera lengua.

fuera de la escuela. El estudio concluye que aunque la lengua oficial sea el inglés, el español es a menudo utilizado en el ámbito familiar.

Sánchez-Saus Laserna (2011) evalúa la influencia del uso de la lengua española que hacen estudiantes extranjeros distinguiendo entre tres categorías: bajo, medio y alto. Los resultados muestran un aumento progresivo del caudal léxico a medida que lo hace el uso de la lengua. Asimismo, constata que aquellos informantes con una mejor competencia lingüística suelen hacer un mayor uso de la lengua, de ahí que también actualicen un mayor número de palabras.

2.8. Conclusiones parciales

La disponibilidad léxica ha evolucionado de forma considerable desde sus orígenes en Francia (Gougenheim *et al.*, 1956, 1964) hasta nuestros días. Según el filósofo griego Aristóteles «El principio es también la parte esencial y primera de donde proviene una cosa». Este capítulo comienza precisamente detallando cómo surgieron los estudios en el ámbito de la disponibilidad a través de la consulta de fuentes originales para comprender en qué consisten y conocer la magnitud de las investigaciones que se comentan *a posteriori*.

Le Français Fondamental es la obra de referencia en el ámbito de la disponibilidad léxica. Tras la caída de la IV República y el fin de la Unión Francesa, los objetivos de este vocabulario se redefinieron y se centraron en: la enseñanza escolar; la difusión de la lengua en los estados que acababan de independizarse y donde el francés gozaba de una situación privilegiada o era la lengua oficial; la respuesta a las necesidades de turistas, estudiantes y trabajadores que se trasladan a un país extranjero. Los investigadores franceses Gougenheim *et al.* (1956, 1964) recopilaron datos de 904 escolares de cuatro departamentos de Francia: Dordogne, Marne, Eure y Vendé. La prueba consistía en hacer que cada informante aportase veinte unidades léxicas relacionadas con un centro de interés dado, hasta un total de dieciséis.

Esa metodología ha sido modificada por Dimitrijević (1969) quien introduce la recogida de datos mediante el sistema de listas abiertas en las que los informantes aportan palabras para cada centro en un tiempo determinado. Los criterios

metodológicos más difundidos en este ámbito surgen a raíz del Proyecto Panhispánico de Disponibilidad Léxica en los años 90. La recogida de datos se realiza conforme al sistema de listas abiertas con una duración de dos minutos para cada centro de interés de los dieciséis presentes en la obra francesa. El objetivo del Panhispánico es elaborar diccionarios de disponibilidad léxica de las zonas de habla hispana para realizar comparaciones dialectales, sociolingüísticas, etnolingüísticas, entre otras.

En el ámbito hispano, los estudios nacieron gracias a la labor incesante de López Morales quien publica en el año 1999 el *Léxico disponible de Puerto Rico*. A continuación, florecen los estudios tanto en Hispanoamérica como en España tomando como base el español como lengua materna. Recientemente, cobran están cobrando mayor importancia los estudios de español como lengua extranjera o segunda lengua en investigaciones llevadas a cabo tanto en países hispanoamericanos como en otros países: Carcedo González (2000c) en Finlandia; Moreno Fernández (2003) en Estados Unidos; González Fernández (2013) en Turquía, López González (2009) en Polonia; Bianca Sandu (2014) en Rumanía; entre otros.

Los edición de los datos de las pruebas de disponibilidad léxica es crucial para los análisis que se pretendan realizar sobre todo si el interés reside en la comparación con otros listados de palabras. En 1998, Samper Padilla establece una serie de pautas y criterios comunes que serán secundados por numerosos estudios, sobre todo aquellos que forman parte del Panhispánico. No obstante, existen investigadores que abogan por incluir todas las palabras producidas por los hablantes para poder conocer realmente cuál es su léxico disponible (Ávila Muñoz, 2006; Samper Hernández, 2006; Ávila Muñoz y Villena Ponsoda, 2010, 2011; etc.).

En cuanto al procesamiento de los datos mediante fórmulas matemáticas, existe un consenso casi generalizado a la hora de calcular el índice de disponibilidad mediante la fórmula de Strassburger y Chávez (1987) y el índice de cohesión mediante la ecuación propuesta por Echeverría (1991). Asimismo, existen otros investigadores, como Ávila Muñoz y Villena Ponsoda, que completan sus análisis con la utilización de otros parámetros como el índice de compatibilidad o el índice de descentralización que marcan la mayor o menor semejanza del léxico propuesto por

un informante con respecto al del resto de hablantes. En cambio, García García (2014) realiza una revisión matemática en la que modifica el índice de disponibilidad, el índice de compatibilidad y el índice de descentralización y permite su cálculo mediante el programa SPSS. Con anterioridad, el cálculo del índice de disponibilidad y la obtención del listado de léxico disponible requería la utilización de algún programa específico, entre los que destacan Lexidisp (Moreno *et al.*, 1995), Dispolex (Bartol Hernández y Hernández Muñoz, 2004) y Dispogen (Echeverría *et al.*, 2005).

El léxico resultante es objeto de numerosos análisis desde disciplinas diferentes. Desde el punto de vista de la etnolingüística, permite conocer la situación concreta de una sociedad en un momento determinado; desde la sociolingüística es posible analizar la incidencia de factores sociales en el léxico a través de determinadas variables sociales; desde la dialectología se pueden comparar las diferentes sintopías de una lengua; desde la lingüística aplicada a la enseñanza de lenguas supone una herramienta útil para la selección del léxico y desde la psicolingüística ofrece la oportunidad de conocer cómo se organizan las palabras en la mente de los hablantes. Aunque esas son las disciplinas más presentes en los estudios de disponibilidad léxica no se descartan nuevas aplicaciones; tanto es así que existen investigaciones que se realizan con fines terapéuticos.

La lingüística aplicada a la enseñanza de lenguas está en menor o mayor medida presente en los estudios de disponibilidad léxica ya que surgieron para dar respuesta a unas necesidades educativas específicas (Gougenheim *et al.*, 1964) y concluyen con la creación de un diccionario o listado de palabras disponibles. Por último, se han recogido algunas investigaciones que muestran la influencia de factores en el aprendizaje de idiomas como la edad de aprendizaje o las estancias en el extranjero que posteriormente serán objeto de análisis en el capítulo de análisis cuantitativos.

CAPÍTULO 3. MARCO METODOLÓGICO

3.1. Introducción

3.2. Naturaleza y método de la investigación

3.3. Selección de la muestra

3.3.1. Tamaño de la muestra

3.3.2. Acceso a los informantes

3.3.3. Requisitos de acceso al Máster

3.3.4. Programa de estudios

3.4. Sistema de recogida de los datos

3.4.1. Encuesta sociológica

3.4.2. Prueba de disponibilidad

3.4.3. Prueba del léxico especializado

3.5. Edición y codificación de los datos

3.5.1. Encuesta sociológica

3.5.2. Prueba de disponibilidad

3.5.2.1. Criterios comunes al Proyecto Panhispánico de Léxico
Disponible

3.5.2.2. Criterios propios

3.5.2.3. Características de la edición en lengua española

3.5.2.4. Características de la edición en lengua extranjera

3.5.3. Prueba del léxico especializado

3.6. Conclusiones parciales

3.1. Introducción

Los interrogantes y objetivos planteados en este estudio son de diversa índole. Abarcan, desde crear un diccionario de disponibilidad léxica, hasta conocer los factores que inciden en una mayor o menor cantidad de vocabulario individual. Para vislumbrar el resultado de las hipótesis planteadas, es necesario aplicar una metodología basada en una serie de procedimientos adecuados. Según Taylor y Bogdan (1987: 15) «el término metodología designa el modo en que enfocamos los problemas y buscamos las respuestas».

El propósito de este capítulo es presentar y delimitar los fundamentos y métodos seguidos para llevar a cabo nuestro estudio. Una de las primeras decisiones en cualquier investigación es la delimitación de la población, entendida como el conjunto de unidades sobre las que se desea conseguir cierta información. Su elección vendrá dictada por los objetivos y el problema de la investigación, los cuales requerirán una ubicación (espacial y temporal) y características determinadas. Una vez definida la población, se procede al diseño de la muestra y a una previa evaluación sobre la accesibilidad a los informantes de los que se pretende recabar los datos.

A continuación, se detalla la elaboración de los tres documentos utilizados para recabar los datos: la encuesta sociológica, la prueba de disponibilidad y la prueba de identificación del léxico especializado. El diseño de la recogida de datos lleva consigo la edición y codificación posterior de los datos. De esta forma, las preguntas del cuestionario y del léxico fueron meticulosamente seleccionadas en función de los objetivos de la investigación pero, también, sobre la interpretación que se pueda realizar del análisis de esos datos hasta llegar a las conclusiones finales.

Por último, se incluyen las pautas de edición seguidas con cada una de las tres herramientas de investigación. El tratamiento de las variables socioculturales tiene como objetivo ser parte integrante de una base de datos manipulable en función de las indagaciones pertinentes. El léxico disponible persigue el cotejo con estudios previos y el análisis de los vocablos resultantes desde diferentes ópticas y que, por tanto, son factores que se tendrán en cuenta en la lematización de las palabras. Las respuestas del léxico especializado se codificarán en forma de escala para que evidencie la capacidad de identificación de tecnicismos de los informantes.

3.2. Naturaleza y método de la investigación

La naturaleza de este trabajo combina una metodología mixta, es decir, cuantitativa y cualitativa. Sendos métodos guiarán el resto de apartados presentes en este capítulo: selección de la muestra, recogida, edición, codificación, procesamiento y presentación de los datos.

La metodología cualitativa cobra gran relevancia con el florecimiento de los estudios relacionados con la sociología y la antropología a partir del siglo XIX. El trabajo de Le Play (1855) sobre familias y comunidades europeas es considerado por Taylor y Bogdan (1987: 17) como la primera investigación científica que utiliza métodos cualitativos de forma consciente. No obstante, ambos autores realizan un recorrido histórico de esta metodología haciendo referencia a la obra de Wax (1971) quien afirma que los métodos empíricos son tan antiguos como la historia escrita. Según Wax, las entrevistas y la observación descriptiva han sido llevadas a cabo por escritores, historiadores y viajeros que van desde el griego Heródoto hasta Marco Polo.

Las técnicas de estudio cualitativas son esencialmente inductivas y narrativas. El investigador diseña un plan de trabajo flexible que va perfilando a medida que recaba información a través de entrevistas y otro tipo de observaciones. Maxwell (2009: 236) define como sigue las estrategias de análisis cualitativas:

A basic principle of qualitative research is that data analysis should be conducted simultaneously with data collection (Coffey & Atkinson, 1996, p. 2). This allows you to progressively focus your interviews and observations, and to decide how to test your emerging conclusions.

Strategies for qualitative analysis fall into three main groups: categorizing strategies (such as coding and thematic analysis), connecting strategies (such as narrative analysis and individual case studies), and memos and displays.

A pesar de la inexistencia de un plan de trabajo estandarizado, la recogida de datos se realiza de forma compleja y rigurosa, lo que hace que la figura del entrevistador sea clave. Por un lado, deberá huir de opiniones e impresiones propias y, por otro lado, tendrá que situarse en lugar de los informantes para conocer mejor la realidad objeto de estudio.

Como apunta Yin (*cf.* 1994: 19)¹³⁴ toda investigación empírica lleva consigo de forma implícita, si no explícita, un diseño. Por tanto, el hecho de no plantear hipótesis constatables o carecer de variables cuantificables no significa que las investigaciones cualitativas no cuenten con un diseño científico. Becker, Geer, Hughes y Strauss (1961: 17) abogan por una definición de diseño mucho más amplia:

If we take the idea of design in a larger and looser sense, using it to identify those elements of order, system, and consistency our procedures did exhibit, our study had a design. We can say what this was by describing our original view of the problem, our theoretical and methodological commitments, and the way these affected our research and were affected by it as we proceeded.

Como contrapartida al empirismo, nace en el primer tercio del siglo XIX el positivismo de la mano de Comte (1896)¹³⁵ y a principios del siglo XX el neopositivismo o positivismo lógico con Durkheim (1938, 1951). Ambos autores son los referentes clásicos de los fundamentos de la metodología cuantitativa. Los procedimientos cuantitativos organizan los datos en variables medibles que corresponden a diferentes segmentos de la población. Dichos datos se analizan siguiendo un método más o menos hermético y uniforme para facilitar la comparación con otros estudios. Para verificar o refutar las hipótesis, prima la objetividad sobre la subjetividad lograda a través de análisis cuantificables y principios más o menos generales y estandarizados.

Los análisis derivados del contraste de las hipótesis permiten determinar el grado de asociación o correlación entre las variables de la muestra. Los resultados educidos pretenden ser extrapolables a la población de estudio llegando a universalizar las hipótesis constatadas. Sin embargo, autores como Popper (1959: 27)¹³⁶ rechazan la generalización de dichos supuestos y advierten del riesgo que supone teorizar con meros datos. En esta misma tendencia, Deutscher (1973: 41) critica el fervor por la confiabilidad y la reproducción en la investigación social:

¹³⁴ «Every type of empirical research has an implicit, if not explicit, research design» Yin (1994: 19).

¹³⁵ Obra póstuma.

¹³⁶ «Now it is far from obvious, from a logical point of view, that we are justified in inferring universal statements from singular ones, no matter how numerous; for any conclusion drawn in this way may always turn out to be false: no matter how many instances of white swans we may have observed, this does not justify the conclusion that all swans are white» (Popper, 1959: 27).

We have, in our pursuit of reliability, been absorbed in measuring the amount of error which results from inconsistency among interviewers or inconsistency among items on our instruments. We concentrate on consistency without much concern with what it is we are being consistent about or whether we are consistently right or wrong. As a consequence we may have been learning a great deal about how to pursue an incorrect course with a maximum of precision. It is not my intent to disparage the importance of reliability per se; it is the obsession with it to which I refer.

Teniendo en cuenta que cada metodología presenta unas ventajas e inconvenientes, en este estudio se ha intentado sacar el máximo partido de cada una de ellas según el objetivo perseguido en cada caso. Existen autores destacados que consideran crucial la relación de complementariedad entre ambas metodologías, eliminando así el debate sobre la prevalencia de una sobre otra (Flick, 2002: 264-265; Ruiz Olabuénaga, 1999: 17). Por tanto, se ha optado por un método u otro según el tipo de interrogante y respuesta planteados, complementando así los análisis realizados.

Hernández Campoy y Almeida (*cf.* 2005: 198) destacan cuatro fases clave del proceso de una investigación sociolingüística cuantitativa en las que interviene la metodología empirista: en el muestreo y la obtención de datos; en la elaboración de variables; en el análisis del comportamiento individual y en el muestreo exploratorio. A pesar de ello, según la perspectiva de análisis, domina el uso de una metodología sobre otra:

- El método empirista ha imperado en la perspectiva psicolingüística donde se han analizado las palabras aportadas por los informantes de forma cualitativa. Por ejemplo, en el Capítulo 6 se ha realizado un análisis de la organización de las palabras en el lexicón mental.
- Las técnicas cuantitativas acuñadas en la disponibilidad léxica sirven de referencia a los análisis léxico-estadísticos tanto en la recolección de datos como en su análisis a través de fórmulas matemáticas para calcular los diversos índices comentados en el Apartado 2.5. Una de las ventajas de este método es la cierta uniformidad que confiere a los datos, permitiendo la comparación con estudios precedentes y futuros.

La combinación de ambos métodos, el cualitativo y cuantitativo, es imprescindible para comprender cuáles son los factores que inciden en el mayor o

menor léxico virtual de los informantes desde un punto de vista sociolingüístico. Se produce una relación simbiótica entre ambos de forma que para realizar análisis cuantificables con diferentes variables son necesarias las relaciones cualitativas entre ellas. Moreno Fernández (2009: 310) establece tres principios generales presentes en toda investigación sociolingüística: el principio de la cuantificación, el principio de la responsabilidad y el principio de la representatividad. En la definición del principio de cuantificación¹³⁷ destaca el papel primordial de los análisis cualitativos para la realización de cualquier análisis cuantitativo.

Por último, cabe destacar que cada método lleva asociado un tipo de análisis y, por tanto, el uso de un programa informático oportuno. De esta forma, la utilidad de los recursos empleados lleva a conseguir los objetivos de investigación perseguidos. Esas herramientas han sido las siguientes¹³⁸:

- Lexidisp. Creada en 1996 en la Universidad de Alcalá, es una aplicación útil para calcular los índices de léxico disponible. Este programa nos ha servido no solo para calcular el índice de disponibilidad léxica, sino también para crear los diccionarios del léxico disponible.
- SPSS (Statistical Product and Service Solutions). Tal y como dice el propio nombre del programa, su objetivo es el análisis estadístico. Se ha empleado para llevar a cabo la mayoría de análisis cuantitativos permitiendo evaluar la relación de más de 300 variables. Asimismo, se ha empleado un programa estadístico que complementa a SPSS, denominado AMOS, para el cálculo de ecuaciones estructurales.
- DispoGrafo. Este software, diseñado por un grupo de investigadores de la Universidad de Concepción (Echeverría *et al.*, 2008), ha sido utilizado para realizar las relaciones semánticas de corte psicolingüístico.

¹³⁷ «Principio de cuantificación [...] Tal y como está formulado este principio no niega en absoluto la posibilidad de establecer relaciones entre variables con medios cualitativos, antes bien la complementa, hasta el punto de que sin éstos no sería posible la aplicación de ningún procedimiento cuantitativo» Moreno Fernández (2009: 310).

¹³⁸ Además de otros programas auxiliares y aplicaciones han contribuido a la edición, codificación, búsqueda de información y presentación, como es el caso de Microsoft Excel, PDF, entre otros.

3.3. Selección de la muestra

La búsqueda de documentación es fundamental para identificar la población de estudio y seleccionar las unidades que constituirán una representación a pequeña escala de la población a la que pertenece. A lo largo de este apartado se fundamentan los tres tipos de decisión que según Sankoff (1974: 21-22) se han de tomar en toda investigación cuantitativa:

- (a) The first is to delineate, at least roughly, the geographic, social, etc., boundaries of the speech community or subcommunity of interest, i.e., to define the sampling universe. (b) The second is to assess the possible relevant geographic, social, and sociolinguistic dimensions of variation within the community, i.e., to construct a stratification for the sample. As the goal of sampling is to tap the existing linguistic variation in a community, it is important for the investigator to consider carefully all types of variation which exist. [...] (c) The third type of decision is to settle on the number of informants and the amount of material to be collected from each, i.e., to fix the sample size.

Para apoyar esas decisiones, en este estudio se ha tomado como referencia el cuadro realizado por Rebagliato, Ruiz y Arranz (*cf.* 1996: 74) que representa los niveles de selección desde que se define la población de forma genérica hasta llegar a los participantes, es decir, quienes efectivamente aportan los datos analizados en cualquier investigación. En lo que sigue, se hace uso de ese modelo para recorrer las fases de la delimitación muestral en este estudio:

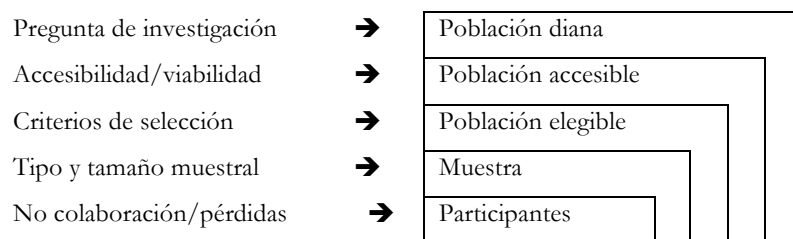


Gráfico 3.1. Niveles de población. Extraída de Rebagliato, Ruiz y Arranz (1996: 74)

- Población diana: conjunto de individuos sobre los cuales se extrapolarán los resultados. La pregunta y respuesta claves para llegar a esa población diana han sido las siguientes: «¿Qué queremos conocer?: El léxico disponible en español y en lengua extranjera del futuro profesorado».
- Población accesible: informantes abordables por el investigador. La investigación se centra en el futuro profesorado formado en la Universidad de Málaga. Para conocer la viabilidad de la recogida de datos de este estudio, se

contactó previamente con el coordinador académico del Máster en Profesorado de Secundaria.

- Población elegible: personas a las que se pretende pasar los cuestionarios y cuya selección va a depender de criterios físicos, económicos, temporales y prácticos, entre otros. En este caso, la población elegible es la misma que la población accesible ya que al haber seleccionado la provincia de Málaga y tratarse de una población finita, es decir, que tiene un número concreto de informantes, se cumplen todos los requisitos necesarios para poder recabar la información de la totalidad de informantes.
- Participantes: tal y como su nombre indica, son los integrantes que forman parte del estudio. Vicisitudes como la falta de asistencia a clase o la imposibilidad de pasar el cuestionario en varias especialidades han hecho menguar la población elegible hasta llegar a los 171 informantes.

Al considerar los informantes como los protagonistas de todo estudio, su selección debe estar bien fundamentada y, como indica Milroy (1987: 46)¹³⁹, ha de estar justificada desde el plano teórico. La elección del futuro profesorado se debe a las necesidades detectadas en el propio ámbito de trabajo, es decir, en enseñanza secundaria. Una de las principales dificultades para implantar la enseñanza bilingüe en los centros andaluces radica en la mejorable competencia lingüística en lengua extranjera del profesorado. En este sentido, la formación inicial de los docentes constituye un pilar básico para la implementación de los cambios necesarios en el gremio. Por ello, el alumnado que cursa el Máster en Secundaria aportará datos de interés para contribuir a la mejora tanto de su formación como la de su futuro alumnado.

3.3.1. Tamaño de la muestra

Una vez seleccionada la muestra de informantes, cabría plantearse cuál es el número mínimo de cuestionarios necesarios para el estudio. Aunque *a priori* la

¹³⁹ «A good judgement sample needs to be based on some kind of defensible theoretical framework; in other words, the researcher needs to be able to demonstrate that his or her judgement is rational and well-motivated» (Milroy, 1987: 46).

respuesta debe ser simple, la cuestión de la idoneidad numérica en una investigación ha sido ampliamente discutida. Si bien existen numerosos estudiosos que han abordado el denominado umbral de significación, no hay ninguna teoría exenta de críticas ni consensuada, ya que como afirman Fox y Bayat (2007: 61) el tamaño de una muestra depende de diferentes consideraciones prácticas como es el tamaño de la población, el muestreo y su representatividad. No obstante, la mayoría de los investigadores, tales como Hernández Campoy y Almeida (2005: 199)¹⁴⁰, coinciden en que resulta inviable en la mayoría de los casos hacer un estudio que englobe a toda la población, de ahí que se seleccione una muestra.

Autores como Labov (1966), aportan datos más concretos sobre esa franja de idoneidad. Afirman que para garantizar la relevancia de los datos, es necesario contar con 25 informantes de una muestra de 100 000, es decir, el 0,0025 % del universo elegido. Sin embargo, Sankoff (1980: 51-52) va más allá proponiendo un número exacto de informantes:

The literature as well as our own experience would suggest that, even for quite complex speech communities, samples of more than about 150 individuals tend to be redundant, bringing increasing data-handling problems with diminishing analytical returns. [...] It is crucial, however, that the sample be well chosen, and representative of all social sub-segments about which one wishes to generalize.

En esa misma línea, Dörnyei y Taguchi (2010: 62)¹⁴¹ comentan la importancia del muestreo para asegurar su representatividad. No parece existir ninguna regla fija que estime el tamaño óptimo de la muestra y ni el tamaño mínimo para que sea representativa. Por ello, sugieren números concretos de informantes según la naturaleza del estudio¹⁴². Asimismo, existen autores que sugieren un número de casos

¹⁴⁰ «Cuando un investigador se propone estudiar la distribución de un determinado rasgo o carácter más o menos relevante entre los miembros de una comunidad (ingresos, altura, gustos literarios, tendencias políticas, etc.) lo normal es que, como hemos visto en IV.4, no estudie a la totalidad de este universo (población), algo que resulta imposible en la mayoría de los casos, sino a una parte, que constituirá la muestra» (Hernández Campoy y Almeida, 2005: 199).

¹⁴¹ Esa misma información se encuentra en la primera edición (Dörnyei, 2003: 74) y en Dörney y Csizér (2012: 82). Se ha seleccionado la segunda edición (Dörnyei y Taguchi, 2010), de ahí que aluda a obras publicadas por el mismo autor en 2007.

¹⁴² «From a purely statistical point of view, a basic requirement is that the sample should have a normal distribution, and a rule of thumb to achieve this, offered by Hatch and Lazaraton (1991), is that the sample should include 30 or more people. [...] From the perspective of statistical significance, the principal concern is to sample enough learners for the expected results to be able to reach statistical significance. [...] a good rule of thumb is that we need around 50 participants to make sure that these coefficients are significant and we do not lose potentially important results.

mínimos según la operación que se pretenda realizar. Etxeberria, Joaristi y Lizasoain (1990: 278) aconsejan utilizar la estadística paramétrica cuando hay más de 10 casos, incidiendo en su efectividad al aumentar los casos, y la estadística no paramétrica con menos de 10 casos. En cuanto al análisis de regresión, Camacho Rosales (2002: 285) establece un mínimo de 10 casos por cada variable independiente para poder calcular el coeficiente de regresión sin sesgos en los valores.

Con respecto a la representatividad de los datos, en la presente investigación no se sigue la distinción realizada por López Morales (1994: 41) entre universo absoluto, es decir, el censo completo, y universo relativo, los candidatos válidos para la investigación. En este caso, el universo absoluto y relativo está compuesto por candidatos representativos y válidos que cumplen el requisito de estar matriculados en el Máster en Profesorado de Secundaria y que los llevará a ser profesores en un futuro. Siguiendo la propuesta de Dörnyei y Csizér (2012: 82), los cuestionarios se pasaron a todo el alumnado del Máster accesible:

When setting the final sample size, it is advisable to leave a decent margin to provide for unforeseen or unplanned circumstances. For example, some participants are likely to drop out of at least some phases of the project; some questionnaires will always have to be disqualified for one reason or another; and – in relation to point 4 above – we may also detect unexpected subgroups that need to be treated separately.

Dado que la población es finita, se solicitó al Centro de Postgrado de la Universidad de Málaga ese censo con los datos referentes a la oferta de especialidades, las solicitudes recibidas y las matrículas realizadas. De esta forma, es

1. In the survey research literature a range of between 1% and 10% of the population is usually mentioned as the ‘magic’ sampling fraction, depending on how careful the selection has been (i.e., the more scientific the sampling procedures applied, the smaller the sample size can be, which is why opinion polls can produce accurate predictions from samples as small as 0.1% of the population).

2. From a purely statistical point of view, a basic requirement is that the results obtained from the sample should have a normal distribution, and a rule of thumb to achieve this, offered by Hatch and Lazaraton (1991), is that the sample should include 30 or more people. However, this is not an absolute rule, because smaller sample sizes can be compensated for by using certain special nonparametric statistical procedures (see Dörnyei, 2007).

3. From the perspective of statistical significance, the principal concern is to sample enough learners for the expected results to be able to reach significance. Because in L2 studies meaningful correlations reported in journal articles have often been as low as 0.30 and 0.40, a good rule of thumb is that we need around 50 participants to make sure that these coefficients are significant and thus we do not lose potentially important results. However, certain multivariate statistical procedures require more than 50 participants; for factor analysis or structural equation modeling, for example, we need a minimum of 100 (but preferably more) subjects» (Dörnyei y Taguchi, 2010: 62).

posible contrastar la población elegible con el número de informantes que completaron las pruebas sociolingüísticas y conocer la representatividad del estudio.

Especialidades	Oferta	Solicitudes	Matrículas	Informantes
Biología y Geología	30	46	26	0
Dibujo, Imagen y Artes Plásticas	20	69	18	11
Economía, Empresa y Comercio	50	63	31	0
Física y Química	20	33	19	0
Formación y Orientación Laboral	20	38	13	7
Ciencias Sociales, Geografía e Historia	45	53	33	17
Hostelería y Turismo	15	36	11	10
Lengua extranjera	65	161	58	38
Lengua, Literatura, Latín y Griego	35	70	33	19
Música	20	25	17	9
Orientación Educativa	60	47	23	10
Procesos Sanitarios	20	23	13	0
Tecnología, Informática y Procesos Industriales	95	115	75	40
Matemáticas	20	63	15	10
Total	515	842	385	171

Tabla 3.1. Plazas del Máster por especialidades y número de informantes

El Máster en Profesorado se divide en catorce especialidades diferentes de las cuales diez han participado en este estudio. En todos los casos, a excepción de la especialidad de Orientación Educativa, la oferta de plazas es menor a las solicitudes. En cambio, las matrículas efectivas descienden considerablemente porque ese año fue el segundo del Máster en Profesorado y muchas solicitudes no fueron aceptadas porque los estudiantes no pudieron acreditar un nivel B1 en un idioma extranjero.

3.3.2. Acceso a los informantes

En esta etapa de la investigación, se hace necesario el contacto con los agentes implicados en la investigación, tanto profesorado como alumnado del Máster en Secundaria, para explicarles la naturaleza del trabajo que se pretendía realizar y pedirles su colaboración. Para llevar a cabo este cometido, se elaboró el documento de negociación¹⁴³ en el que se incluyen datos de interés, tales como: tipo de

¹⁴³ Anexo III.

investigación, duración y tipo de cuestionarios necesarios, confidencialidad de los datos, cometido de la investigación y los datos de contacto de la doctoranda.

A continuación, se estableció comunicación con los coordinadores académicos del máster, quienes expresaron su interés y aceptación. Tras obtener su aprobación, se les envió el documento de negociación y el cuestionario al profesorado que impartía clases en el máster para que concediesen el tiempo de clase necesario para obtener los datos. Una vez conseguido el consentimiento del profesorado, se les mostró el documento de negociación inicial a los representantes del colectivo de estudiantes de las asignaturas en las que se desarrollaría el cuestionario.

Aunque el universo de la investigación elegido era todo el colectivo de estudiantes del Máster en Profesorado, en la práctica no pudo hacerse factible por diversos motivos. Por un lado, había profesores que tenían todas las clases ya planificadas por lo que era inviable pasar los cuestionarios. Por otro lado, el número de informantes no se corresponde con el alumnado matriculado por factores externos a la investigación, tales como:

- Retrasos considerables de los estudiantes que provocaron que no fuese viable la realización de la encuesta.
- Ausencia del alumnado el día que se pasó el cuestionario.
- Motivos personales que fuerzan al alumnado a abandonar la clases antes, por lo que no tuvieron tiempo suficiente para acabar de completar la prueba.
- Falta de interés por colaborar con el estudio.

A pesar de las dificultades encontradas, se recogieron datos de más de la mitad del alumnado matriculado en las especialidades objeto de estudio durante el curso académico 2010/2011, a excepción de la Orientación Educativa (43,48 %). Teniendo en cuenta lo tratado sobre el número de informantes y la calidad de los datos recogidos, este estudio cumple con los criterios de suficiencia y adecuación expuestos por Rodríguez Gómez, Gil Flores y García Jiménez (1996: 75):

Durante esta etapa es preciso *asegurar el rigor* de la investigación. Para ello debemos tener en cuenta los criterios de suficiencia y adecuación de los datos. La suficiencia se refiere a la cantidad de datos recogidos, antes que al número de sujetos. La suficiencia se consigue cuando se llega a un estado de «saturación informativa» y la nueva información no aporta nada nuevo. La adecuación se refiere a la selección de la información de acuerdo con las necesidades teóricas del estudio y del modelo emergente.

En cuanto a los criterios éticos presentes en la investigación, se han tenido en cuenta factores clave como la confidencialidad de los datos, el anonimato y la transparencia a la hora de recoger y presentar los resultados. Los informantes no han escrito su nombre por lo que los cuestionarios han sido numerados según el orden de recogida. En este sentido, cabe mencionar una cita de la obra denominada *Evaluación institucional*, cuyos principios básicos se han respetado en este estudio: «Toda política de evaluación institucional debe sustentarse en la transparencia, la precisión, la comunicabilidad, el rigor y la actualidad de la información, apelando a formas múltiples y avanzadas de transmisión de la misma» (Universidad Nacional del Litoral, 2008: 13).

3.3.3. Requisitos de acceso al Máster Universitario en Profesorado

Se establece como condición general común al resto de másteres oficiales estar en posesión de un Título Universitario Español (Graduado, Diplomado, Ingeniero Técnico, Arquitecto Técnico, Licenciado, Arquitecto o Ingeniero). Los titulados en un país extranjero perteneciente al Espacio Europeo de Educación Superior (EEES) deben estar en posesión de un título expedido por una institución superior que faculte el acceso del máster oficial. Sin embargo, en el caso de los títulos no pertenecientes al EEES, la Universidad comprueba que, en el país expedidor del título, poseen una formación equivalente a los títulos universitarios españoles y no es necesaria su homologación.

En lo que concierne a los requisitos específicos, el perfil profesionalizador del Máster en Profesorado hace necesarias las siguientes condiciones de acceso:

- 1) Poseer las competencias relativas a la especialización que se desee cursar en el máster. La acreditación puede realizarse aportando la titulación universitaria que se corresponda con la especialidad elegida o mediante la superación de una prueba diseñada al efecto por las universidades basadas en el temario de

oposición de la especialidad en cuestión. Aunque en las universidades andaluzas se establece como criterio de admisión principal al Máster en Profesorado la nota del expediente académico, en caso de empate la titulación aportada por los solicitantes se considera el segundo criterio. Para tal efecto, se tienen en cuenta los grupos de preferencias de titulaciones revisados anualmente por la Comisión de Distrito Único Andaluz.

- 2) Acreditar el dominio de una lengua extranjera equivalente al nivel B1 del Marco Común Europeo de Referencia para las Lenguas (MCERL) de acuerdo con la Recomendación N.º R (98)6 del Comité de Ministros de Estados Miembros de 17 de octubre de 2000. Tal verificación de competencia lingüística se realiza siguiendo las directrices fijadas en el Convenio de Colaboración para la Acreditación de Lenguas Extranjeras, suscrito por las Universidades Andaluzas el 2 de julio de 2011. Entre los objetivos y características generales del Convenio, cabe citar los siguientes:

- Validar ciertos certificados de dominio de lenguas extranjeras para la acreditación de los niveles del MCERL. Para ello, en el Convenio se adjuntan los certificados más conocidos clasificados según el nivel del MCERL de forma que todos aquellos que correspondan a un nivel B1 o superior serán admitidos como válidos.
- Reconocer el nivel B1 de dominio de lenguas extranjeras a aquellos titulados que acrediten haber superado en su formación universitaria al menos 24 créditos en una lengua extranjera de la Unión Europea.
- Organizar exámenes de acuerdo con las directrices del MCERL para aquellos estudiantes que no cuenten con ningún certificado. De esta forma, la acreditación se producirá por medio de un examen al margen de que el estudiante haya realizado algún curso de idiomas en la universidad o por otra vía.
- Planificar debidamente las pruebas de idiomas teniendo en cuenta las lenguas, los niveles y las convocatorias que se produzcan.

3.3.4. Programa de estudios

Al tratarse de un máster oficial, el plan de estudios se ha organizado en módulos de un total de 60 créditos europeos. La fundamentación normativa sobre los contenidos se encuentra en la Orden ECI/3858/2007, de 27 de diciembre, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habilite para el ejercicio de las profesiones de profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas.

- a. Módulo genérico común a las distintas especialidades del máster que consta de un total de 12 créditos.
- b. Módulo específico de 24 créditos con asignaturas de la especialidad y la asignatura de Innovación Docente e Iniciación a la Investigación en su Área.
- c. Prácticum de 16 créditos distribuidos en prácticas externas y el trabajo fin de máster.
- d. Para completar la formación, se añaden 8 créditos de libre designación. El estudiante podrá conseguir esos créditos cursando materias/asignaturas de otras especialidades del Máster en Profesorado o de otros másteres oficiales, materias ofertadas por cada Universidad y asistir a conferencias.

Como se puede apreciar en el módulo específico, los títulos se han configurado para que el estudiante alcance una cierta especialización, denominada itinerario curricular, mención o en el caso de la Universidad de Málaga, especialidad. Estas menciones que rigen el módulo específico se han estructurado según las especialidades docentes de los cuerpos de profesores de Enseñanza Secundaria y los cuerpos de profesores técnicos de Formación Profesional.

No obstante, cada universidad oferta determinadas especialidades o un máster de carácter generalista. Para poder tener derecho a presentarse a las oposiciones de Secundaria es necesario acreditar la superación del Máster en Profesorado, independientemente de la especialidad realizada. La Universidad de Málaga cuenta con una amplia oferta de especialidades del Máster en Profesorado que se distribuyen

en horario de mañana y tarde. A continuación, se expone el plan de estudios del curso 2010/2011¹⁴⁴:

a. Módulo general

Especialidad	Asignatura	Créditos
Todas	Aprendizaje y desarrollo de la personalidad	4
	Procesos y contextos educativos	4
	Sociedad, familia y educación	4

Tabla 3.2. Asignaturas del módulo general del Máster en Profesorado

b. Módulo específico: asignatura de 6 créditos adaptada a la especialidad del máster llamada Innovación docente e iniciación a la investigación educativa (área Ciencia y Tecnología, Filología, etc.).

Área	Especialidad	Créditos
Ciencia y Tecnología	1. Biología y Geología	6
	2. Física y Química	6
	3. Matemáticas	6
	4. Tecnología, Informática y Procesos Industriales (TIPI)	6
Filología	5. Lengua extranjera: Francés	6
	6. Lengua extranjera: Inglés	6
	7. Lengua y Literatura: Latín y Griego	6
Ciencias Sociales	8. Geografía e Historia	6
	9. Orientación Educativa	6
Formación Profesional	10. Formación y Orientación Laboral	6
	11. Economía, Empresa y Comercio	6
	12. Procesos Sanitarios	6
Arte y Deporte	13. Música	6
	14. Educación Física	6

Tabla 3.3. Áreas y especialidades del Máster en Profesorado

La asignatura titulada Diseño y desarrollo de programaciones y actividades formativas se imparte en todas las especialidades pero es específica para cada una de ellas. Cuenta con 8 créditos en todas las especialidades excepto en Orientación Educativa que le dedica 4 créditos.

Especialidad	Asignatura	Créditos
1	Complementos para la formación disciplinar: Biología	3
	Complementos para la formación disciplinar: Geología	3
	El Currículum de Biología y Geología en ESO y Bachillerato	4
2	Complementos para la formación disciplinar: Física	3

¹⁴⁴ Es el curso donde se recogieron los cuestionarios.

	Complementos para la formación disciplinar: Química	3
	El Currículum de Física y Química en ESO y Bachillerato	4
3	Pensamiento matemático, historia de las matemáticas y resolución de problemas	6
	El Currículum de Matemáticas en ESO y Bachillerato	6
4	Complementos para la formación disciplinar: Informática	3
	Complementos para la formación disciplinar: Tecnología y procesos industriales	3
	El Currículum de Tecnología, Informática y Procesos Industriales	4
5	Complementos para la formación disciplinar: Inglés	6
	El Currículum de Inglés en ESO y Bachillerato	4
6	Complementos para la formación disciplinar: Francés	6
	El Currículum de Francés en ESO y Bachillerato	4
7	Complementos para la formación disciplinar: Lengua y Literatura. Latín y Griego	6
	El Currículum de Lengua y Literatura, Latín y Griego en ESO y Bachillerato	4
8	Complementos para la formación disciplinar: Geografía e Historia y Filosofía	6
	El Currículum de Geografía e Historia y Filosofía en ESO y Bachillerato	4
9	Complementos para la formación disciplinar: Orientación Educativa	6
	Diseño, desarrollo y evaluación de un currículum basado en competencias	4
	Orientación e intervención socioeducativa	4
10	La Formación Profesional y el empleo	3
	Modelos de Formación y Orientación	3
	El Currículum de Formación y Orientación Laboral en la Formación profesional	4
11	Complementos para la formación disciplinar en Economía.	3
	Complementos para la formación disciplinar en Organización y Administración de Empresas	3
	El Currículum de Economía, Empresa y Comercio en la Formación profesional	4
12	Educación para la salud	3
	Organización de las profesiones sanitarias	3
	El Currículum de Procesos Sanitarios en la Formación profesional	4
13	Complementos para la formación disciplinar: Música	6
	El Currículum de Música	4
14	Complementos para la formación disciplinar: Dibujo, Imagen y Artes Plásticas	6
	El Currículum de Dibujo, Imagen y Artes Plásticas	4

Tabla 3.4. Asignaturas del módulo específico del Máster en Profesorado

c. Prácticum:

Especialidad	Asignatura	Créditos
Todas las especialidades	Prácticas externas	10
	Trabajo fin de máster	6

Tabla 3.5. Asignaturas del prácticum del Máster en Profesorado

d. Libre configuración:

Especialidad	Asignatura	Créditos
Todas las especialidades	Conferencias: La profesión docente	8

Tabla 3.6. Asignaturas de libre configuración del Máster en Profesorado

3.4. Sistema de recogida de los datos

Con objeto de hacer viable y efectiva la recogida de datos, se han tenido en cuenta una serie de criterios a la hora de elaborar las estrategias o herramientas de investigación. Como apunta Moreno Fernández (2009: 313):

Las técnicas o estrategias para conseguir materiales lingüísticos se pueden clasificar a lo largo de un continuum que representa distintos grados de estructuración. La estructuración determina la previsibilidad de los materiales que se van a reunir: cuanto más estructurada sea la técnica, con más seguridad se conocen los posibles materiales que se van a producir; cuanto más estructurada, más fácil es proceder al análisis cuantitativo de los materiales. Cuando se aplica una técnica poco estructurada no es fácil prever qué materiales se van a reunir, por lo que su análisis cualitativo se hace más complicado y, como consecuencia de ello, también la cuantificación.

A grandes rasgos, las consignas de este estudio han sido las siguientes:

- La extensión de la encuesta debía ser lo suficientemente extensa para dar respuesta a las hipótesis de investigación, pero adecuada para evitar la desidia por parte de los informantes a la hora de responderla.
- El espacio físico y temporal de recogida de los cuestionarios tenía que comprender un periodo inferior a una hora y media para evitar que posibles contratiempos impidieran finalizar la prueba en las clases de esa duración.
- La estructura de realización de la prueba precisaba estar bien definida de forma que en todas las sesiones se pudiese llevar a cabo de la misma manera.
- El diseño de las preguntas tenía que ser minuciosamente estudiado para que diese respuesta a las hipótesis de investigación.
- Cada variable debía ser fácilmente codificable para que pudiese formar parte de la base de datos SPSS.

En cuanto al tipo de cuestionarios, Hernández Campoy y Almeida (2005: 116) distinguen entre directos e indirectos; en otras palabras, presenciales y no presenciales. En este caso, se trata de cuestionarios directos puesto que han sido pasados en papel permitiendo la interacción del investigador con los informantes, tan necesaria para establecer aclaraciones pertinentes sobre la recogida de datos. Sin embargo, Moreno Fernández (1990: 194) clasificaría este *modus operandi* como técnica indirecta ya que el hablante proporciona la información por sí mismo de forma inconsciente.

A continuación, se exponen las partes del cuestionario empleado y las fuentes en las que se basa su diseño ya que algunos aspectos se han adaptado a las necesidades del estudio. Además, se han seguido las recomendaciones dadas por De Gortari (1979: 18)¹⁴⁵ y se han cuidado con detalle las técnicas de recogida de datos y su posterior codificación por considerarlas una parte esencial del trabajo de investigación.

3.4.1. Encuesta sociológica

Los estudios tradicionales de disponibilidad léxica incluyen, además de la prueba del léxico, preguntas de carácter sociológico. En el proyecto Lexidisp (Moreno Fernández, Moreno Fernández y García de las Heras, 1995) se incorporan las siguientes cinco variables: 1) sexo (hombre, mujer); 2) centro escolar (público, privado); 3) comunidad (rural, urbana); nivel sociocultural (alto, medio, bajo); edad (1.^a, 2.^a, 3.^a y 4.^a generación).

The screenshot shows a window titled "Selección" with the following elements:

- Buttons: "aceptar" and "cancelar" at the top right.
- Checkbox: ☐ Todas
- Variable 1: ☒ Var 1 (sexo). Dropdown menu showing "1. (hombre)" and "2. (mujer)".
- Variable 2: ☒ Var 2 (centro escolar). Dropdown menu showing "1. (publico)".
- Variable 3: ☒ Var 3 (comunidad). Dropdown menu showing "2. (rural)".
- Variable 4: ☐ Var 4 (nivel sociocultural). Dropdown menu showing "1. (alto)".
- Variable 5: ☐ Var 5 (edad). Dropdown menu showing "1. (1ª generación)".

Figura 3.1. Ejemplo de variables sociológicas en Lexidisp

No obstante, el investigador puede definir otra variable según las características del estudio otorgándoles otros valores. Por ejemplo, en este caso se ha

¹⁴⁵ «Las técnicas se inventan, se enseñan, se aprenden, se transmiten de manera oral o escrita y a través de su demostración en la actividad misma. Su investigación y su perfeccionamiento son siempre obra colectiva y progresiva. Las técnicas se desarrollan continuamente, se multiplican sin cesar y cambian al paso y la medida en que avanzan la ciencia y la tecnología. En todo caso las técnicas constituyen una de las partes más adelantadas de la actividad científica» (De Gortari, 1979: 18).

modificado la cuarta variable por la lengua seleccionada para completar el cuestionario. La variante 1 corresponde al idioma inglés, en lugar de al nivel sociocultural alto, mientras que la 2 al francés, en vez de al nivel sociocultural medio. Se puede ver en la figura que el formato está configurado de tal forma que solo se pueden elegir unas variables determinadas. Por ello, los cambios realizados no se visualizan en el programa pero son intercambiables por el usuario que lo tendrá en cuenta a la hora de interpretar los resultados.

Los datos se han introducido en el programa mediante un archivo de texto por cada uno de los centros de interés. Los análisis se realizaron por idiomas para poder confeccionar un diccionario con el léxico disponible de forma independiente. De esta forma, se crearon nueve documentos por idioma relativos a cada centro de interés, sumando un total de veintisiete archivos. El formato de los documentos viene bien definido en el manual de usuario de Lexidisp, el cual se explica a partir de este ejemplo:

11211 122 01 cabeza, mano, pie, uña, dedo, diente, pelo, espalda, corazón, meñique, pie, pierna, brazo, cuello

- Los cinco primeros dígitos corresponden a las variables sociológicas. Están representadas con un valor numérico de 1 a 4, siendo las posibles opciones que puede tener cada variable. En el caso expuesto en el ejemplo y según la codificación predefinida de Lexidisp, el informante sería un hombre de un centro escolar público, de una comunidad rural, con un nivel sociocultural alto y perteneciente al primer rango de edad.
- El número de informante se escribe a continuación separado por un espacio. En este ejemplo se trata del informante 122. A cada sujeto le corresponde una línea con resultados en el archivo, por lo que en lengua española el documento contendrá 171 líneas correspondientes a los 171 informantes, en inglés 150 líneas y en francés 21 que equivale a los estudiantes que han realizado el cuestionario en esos idiomas.
- El número que identifica al centro de interés está formado siempre por dos dígitos y va igualmente precedido de un espacio. En el ejemplo, se observa

que se ha puesto un cero antes del número del centro, en este caso el primer centro referido a las partes del cuerpo humano. En el caso de centros de interés o de informantes con una sola cifra, irá precedido de un cero.

- Los vocablos aportados por cada informante en cada caso se sitúan tras los datos numéricos separándose de ellos por un espacio en blanco. Cada vocablo va seguido de una coma y un espacio, excepto la última palabra de cada línea, es decir, de cada informante. La coma es un factor decisivo a la hora de separar las palabras y vocablos, ya que si no hay coma y aparece *cabeza mano* el programa lo procesa como un solo vocablo y se contabiliza posteriormente como una unidad léxica en lugar de dos. Cualquier error de este tipo es fácilmente localizable cuando se extrae la lista de palabras según el índice de disponibilidad léxica ya que aparecería una entrada con escasa disponibilidad, *cabeza mano*.

Además de las cinco variables introducidas en Lexidisp, se han recogido variables socioculturales que han permitido un análisis sociolingüístico más específico con el programa estadístico SPSS. En este sentido, cabe destacar la investigación de Ávila Muñoz y Villena Ponsoda (2010) que incorpora un gran abanico de variables extralingüísticas.

Como se puede observar en el Anexo IV, la encuesta comienza con un apartado dedicado a recabar datos identificativos. Sirven para conocer el grupo, el idioma en el que se realiza la prueba y la fecha exacta en que se completó el cuestionario. Una vez recogido, se le añade un número al cuestionario según el orden de recogida para localizarlo de forma rápida en caso necesario.

A continuación, siguen 27 preguntas sociológicas y reticulares para conocer cuáles son los datos relacionados con la formación y con el contacto con otras lenguas tanto del informante como de sus familiares. Asimismo, se incluyen cuestiones relacionadas con el nivel económico del informante, con su movilidad geográfica y con la densidad de sus relaciones. Algunas de esas preguntas, han sido extraídas de la investigación llevada a cabo en Málaga por Ávila Muñoz y Villena Ponsoda (2010: 313-314). Tras el apartado de cuestiones sociológicas, se incluyeron un total de diez folios en blanco destinados a la prueba de léxico disponible.

3.4.2. Prueba de disponibilidad

Las pruebas de disponibilidad léxica realizadas en el Français Fondamental han experimentado diferentes modificaciones a lo largo del tiempo, sobre todo en lo referente a los centros de interés y a la metodología de recogida del léxico. Originariamente, el vocabulario se recogía mediante listas cerradas en las cuales los informantes debían escribir veinte palabras en cada centro de interés sin límite de tiempo (Gougenheim *et al.*, 1964)¹⁴⁶.

El Proyecto Panhispánico de Disponibilidad Léxica (PPHDL) recoge los dieciséis centros de interés presentes en el Français fondamental con ligeros matices en la denominación, por ejemplo *El cuerpo humano* es denominado *Partes del cuerpo humano*. Una de las bases del proyecto reside precisamente en la utilización de los mismos centros y metodología, por lo que algunos autores han ampliado los centros manteniendo los dieciséis ya establecidos.

Uno de los centros nuevos más utilizados es *Los colores* presentes en las investigaciones de Justo Hernández (1986) en México, Butrón (1987) en Puerto Rico, Benítez Pérez (1992) en Madrid, Etxebarria Arostegui (1996) en el País Vasco, Samper Padilla y Hernández Cabrera (1997) en Gran Canaria, Mateo García (1998) en Almería, Gómez Devís (2004) en Valencia, Casanova Ávalos (2006) en Castellón, Ayora Esteban (2006) en Ceuta, Ahumada Lara (2006) en Jaén, Arnal Purroy (2008) en Aragón, Ávila Muñoz y Villena Ponsoda (2010) en Málaga. Prado Aragonés y Galloso Camacho (2005) incorporan también en Huelva el centro de *Los colores* en la mayoría de sus cuestionarios exceptuando la zona de Moguer donde introdujeron *La inmigración* y en Ayamonte *Portugal e inmigración*. Además, se incorporó otro centro de interés poco común como *la salud*.

Sin embargo, en las investigaciones que no forman parte del Proyecto Panhispánico la elección de centros de interés es mucho más laxa. Por ejemplo, Borrego Nieto y Fernández Juncal (2002) realizaron un estudio con solo diez centros para conocer el léxico disponible de los jubilados cántabros, con temas nuevos como *Agricultura y ganadería* y *La pesca*. Del mismo modo, Vargas Sandoval (1991) trabaja

¹⁴⁶ Vid. apartado 2.2.2 para una revisión de los principios epistemológicos de la disponibilidad léxica.

con seis centros nuevos: 1) *Arte*; 2) *Derecho*; 3) *Economía*; 4) *Política*; 5) *Sentimiento*; 6) *Teología*.

Teniendo en cuenta que este estudio se realiza en dos lenguas diferentes, la realización de la prueba en cada idioma sobre los dieciséis centros tradicionales supondría un total de 64 minutos efectivos solo para realizar la prueba del léxico disponible. Por tanto, el número se redujo a nueve en cada lengua, con lo que la prueba del léxico duraría un mínimo de 36 minutos. Los criterios de selección del léxico han sido los siguientes:

- Centros de interés presentes en los estudios clásicos que permitiesen realizar comparaciones posteriores.
- Temas relacionados con el nivel B1 del Marco Común Europeo de Referencia para las Lenguas. El MCERL dedica un apartado referido a la competencia léxica donde se exponen dos tablas que ilustran el conocimiento y la capacidad de controlar el vocabulario según los seis niveles de referencia. El nivel mínimo exigido a los estudiantes es el B1, donde la escala especifica:

Riqueza de vocabulario: Tiene suficiente vocabulario para expresarse con algún circunloquio sobre la mayoría de los temas pertinentes para su vida diaria como, por ejemplo, familia, aficiones e intereses, trabajo, viajes y hechos de actualidad.

Dominio del vocabulario: Manifiesta un buen dominio del vocabulario elemental, pero todavía comete errores importantes cuando expresa pensamientos más complejos, o cuando aborda temas y situaciones poco frecuentes (Consejo de Europa, 2002: 109).

González Fernández (2014) evalúa la idoneidad de los centros de interés clásicos en una investigación en Turquía sobre el léxico disponible en español como lengua extranjera. A través de un análisis cuantitativo sobre la producción léxica en cada centro de interés, deja entrever cuáles son los centros que deberían ser modificados por otros más productivos. Concretamente se refiere a los centros con menor número de palabras en estudios previos (*Trabajos del campo y el jardín*, *El campo*, *La calefacción y la iluminación* y *Los objetos colocados a la mesa para la comida*) y los centros que más de un 30 % de los informantes han dejado en blanco (*El campo*, *Los trabajos del campo y el jardín*, *La ciudad*, *Los juegos y las distracciones*, *La calefacción y la iluminación* y *Los objetos colocados a la mesa para la comida*).

- Vocabulario relacionado con la práctica docente del profesorado. Entre las posibilidades de elección que ofrece el PPHDL y los temas del Consejo de Europa, se han seleccionado aquellos más relevantes en el ámbito educativo, tanto por su presencia en los libros de texto como por su utilización en clase.

Español	English	Français
01. Partes del cuerpo	01. Human body parts	01. Parties du corps humain
02. La ropa	02. Clothes	02. Les vêtements
03. Comidas y bebidas	03. Food and drink	03. La nourriture et les boissons
04. La escuela (muebles y materiales)	04. The school (furniture and school material)	04. L'école (meubles et matériel)
05. La ciudad	05. The city	05. La ville
06. Medios de transporte	06. Means of transport	06. Moyens de transport
07. Juegos y distracciones	07. Games and entertainments	07. Jeux et loisirs
08. Profesiones y oficios	08. Professions and Jobs	08. Professions et métiers
09. Ordenadores e internet	09. Computer and Internet	09. Ordinateurs et Internet

Tabla 3.7. Centros de interés en los tres idiomas del estudio

El último centro de interés, *Ordenadores e internet*, no cumple el primer requisito pero sí los dos siguientes. El segundo criterio se satisface por tratarse de un tema relacionado con la actualidad y la vida diaria de los estudiantes, ya que los ordenadores e internet forman parte de nuestras actividades diarias. En cuanto al tercer criterio, su justificación se debe a la importancia en la enseñanza actual de las nuevas tecnologías de la información.

Siguiendo la metodología del PPHDL, la realización de la prueba se realizó mediante el sistema propuesto de listas abiertas en la que los informantes tenían que escribir palabras relacionadas con el tema durante dos minutos. A diferencia de algunas investigaciones (Samper Padilla *et al.*, 2003, entre otras), los folios en blanco situados tras el cuestionario sociológico no contienen ningún epígrafe que indicase el centro de interés. Por tanto, en cada clase se decía en voz alta el centro de interés y se escribía en la pizarra en el momento dando paso, a continuación, a los dos minutos de escritura. De esta forma, el informante solo reflexiona sobre el centro y sus posibles asociaciones en el momento de completar el cuestionario y no puede omitir ninguno.

Con objeto de que las listas fuesen realmente abiertas y los estudiantes no viesen limitada la producción léxica, en el cuadernillo entregado a los informantes se adjuntaron diez folios en blanco de manera que ellos mismos podían distribuir las

palabras en las columnas que necesitasen. De esta forma, no es necesario dar las instrucciones que, según Hernández Muñoz (2006: 262), requieren las investigaciones con columnas predefinidas. En esos casos, se sugiere a los informantes que escriban al dorso una vez que agoten el espacio reservado para el léxico disponible en las columnas.

3.4.3. Prueba del léxico especializado

Para completar los resultados del léxico disponible, se ha realizado una prueba específica centrada en la identificación del léxico especializado que toma como referencia la investigación previa llevada a cabo por Ávila Muñoz y Villena Ponsoda (2010). Según los autores (Ávila Muñoz y Villena Ponsoda, 2010: 182), los resultados de este test se traducen en una medida denominada Pident¹⁴⁷ para «contrastar de manera independiente la relación comprobada en la correlación bilateral entre la amplitud léxica (IAL) y la descentralización (IDD) del vocabulario de un hablante».

Por un lado, la investigación de Ávila Muñoz y Villena Ponsoda (2010) presenta resultados relevantes que muestran una correlación positiva con un nivel de significación bilateral $<0,01$ entre la amplitud léxica y el léxico especializado. Por otro lado, establece asociaciones entre el Pident y el nivel educativo de los informantes de tal forma que a mayor nivel de instrucción formal, mejor será el conocimiento del vocabulario especializado. Según Ávila Muñoz y Villena Ponsoda (2010: 218):

La relación de Pident con la estratificación educacional de la comunidad de habla es más fuerte que la que se establece entre esta y la descentralización léxica (ICoDD). [...] El conocimiento especializado parece menos accesible a los hablantes poco instruidos, en tanto que estos tienen un relativamente mayor acceso al léxico general. Del mismo modo, los hablantes con instrucción media se igualan a los universitarios en el vocabulario general pero no llegan, con mucho, a la competencia léxica de estos en el conocimiento y manejo del vocabulario especializado.

En este estudio, la prueba del léxico especializado se ha confeccionado en español, inglés y francés. Al igual que en la prueba de disponibilidad léxica, cada informante tenía que responder a una prueba de identificación del léxico especializado en español y a otra en la lengua extranjera seleccionada. El test consiste en asignar a dieciocho palabras especializadas el número del centro de interés

¹⁴⁷ La descripción precisa de los diferentes índices matemáticos aplicados se encuentra en el apartado 2.5.4.

correspondiente de entre los nueve presentes en la prueba de disponibilidad léxica. Por lo tanto, cada informante debía clasificar dos palabras por centro de interés¹⁴⁸.

3.5. Edición y codificación de los datos

Como se ha definido al principio del capítulo, este estudio analiza de forma cualitativa y cuantitativa el léxico disponible del futuro profesorado, optando por una metodología mixta. Maxwell (1996: 107) marca la diferencia en la codificación de los datos según la metodología aplicada:

The main categorizing strategy in quantitative research is coding. This is rather different from coding in qualitative research, which consists of applying a pre-established set of categories to the data according to explicit, unambiguous rules, with the primary goal being to generate frequency counts of the items in each category. In qualitative research, in contrast, the goal of coding is not to produce counts of things but to «fracture» (Strauss, 1987, p. 29) the data and rearrange it into categories that facilitate comparison between things in the same category and between categories.

La categorización cuantitativa está presente en las tres partes del cuestionario, lo cual permite traducir cada respuesta al lenguaje numérico. Por una parte, la encuesta sociológica cuenta con una mayoría de preguntas cerradas para facilitar su interpretación. Por otra parte, las respuestas de la prueba de disponibilidad léxica y de identificación del léxico especializado han sido contabilizadas teniendo en cuenta en la primera el número de palabras producidas y sus diferentes características, tales como las faltas de ortografía, y en el segundo caso el número de respuestas correctas¹⁴⁹, incorrectas y no contestadas.

La interpretación cualitativa se ha llevado a cabo estableciendo categorías definidas antes y después de la recogida de datos, ya que el propio análisis ha demandado la fractura mencionada por Strauss (1987: 29). Ejemplo de ello ocurre con las anotaciones y clasificaciones realizadas del léxico disponible. En un primer momento, el análisis de algunas categorías, como las faltas de ortografía, no se consideró relevante debido al elevado nivel académico de los informantes. Sin embargo, a medida que se lematizaban los datos su análisis se vislumbraba oportuno.

¹⁴⁸ *Vid.* Anexo IV para consultar la prueba realizada.

¹⁴⁹ Una respuesta se considera correcta cuando el número asignado a la palabra especializada corresponde, en efecto, a su centro de interés. En caso contrario, se considerará incorrecta.

3.5.1. Encuesta sociológica

El plan de codificación establecido para los datos sociológicos atribuye un valor numérico a cada respuesta. A continuación, esta pasa a formar parte de la base de datos en el programa SPSS en forma de variable independiente. El objetivo es hacer cuantificables los resultados de la prueba para poder extraer conclusiones. En esta línea, cabe citar la definición de medición aportada por Stevens (1946: 1):

But measurement is a relative matter. It varies in kind and degree, in type and precision. In the broadest sense measurement is the assignment of numerals to objects or events according to rules. And the fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement. The rules themselves relate in part to the concrete empirical operations of our experimental procedures which, by their sundry degrees of precision, help to determinate how snug is the fit between the mathematical model and what it stands for.

Una vez cuantificadas las respuestas, el siguiente paso consistió en diferenciar la naturaleza de la variable que condiciona el tipo de análisis estadísticos que se puede llevar a cabo con cada una de ellas. Para la mencionada elección, se toma como referencia la clasificación de niveles propuestas por Salkind (1998).

- «El nivel nominal de medición, de la palabra latina *nomin* (nombre), describe variables de naturaleza categórica y que difieren en calidad más que en cantidad» (Salkind, 1998: 113-114). Por ejemplo, en la encuesta se ha codificado con el número 1 a los informantes que han contestado en inglés y con el número 2 aquellos que han realizado el cuestionario en francés. La asignación ha sido arbitraria y privativa puesto que ni el número 1 ni el 2 definen *per se* el idioma y ningún informante puede cumplir ambas opciones ni situarse en un intervalo.
- «El nivel de intervalo de medición, del latín *interval lum* (que significa espacios entre paredes) describe variables que tienen intervalos iguales entre ellas» (Salkind, 1998: 115). Esta variable aparece en el programa estadístico denominada «Escala» y permite incluir datos que forman parte de un continuo. Este es el caso de los índices de medición léxica: el índice corregido de descentralización léxica, el índice corregido de amplitud léxica, el índice corregido de identificación de tecnicismos y el índice de disponibilidad léxica,

entre otros. De esta forma, no se producen sesgos en los datos ya que se presentan las cifras concretas que equivalen a las respuestas de los informantes según la fórmula matemática aplicada en cada caso.

- «El nivel ordinal de medición describe variables que se pueden ordenar a lo largo de algún tipo de continuo» (Salkind, 1998: 114). Para poder establecer comparaciones entre grupos, en ocasiones variables de intervalo han sido codificadas como variables ordinales. Por ejemplo, se organizaron los resultados de las pruebas del léxico especializado y se establecieron cuatro niveles que van desde 0 a 100. Estos rangos permitieron establecer un orden entre las respuestas de manera que el primer intervalo correspondería a un número de respuestas inadecuado hasta llegar al cuarto nivel considerado adecuado. Además, esta clasificación permitió la correlación de variables nominales, como es el sexo, con el conocimiento del léxico especializado.
- «El nivel de razón de medición, del latín *ratio* (que significa cálculo), describe variables que tienen intervalos iguales entre ellas pero que también tienen un cero absoluto» (Salkind, 1998: 116). En este estudio no se ha codificado ninguna variable de este tipo.

Cabe destacar que para la asignación numérica de cada respuesta de la variable se han tenido en cuenta diferentes factores e investigaciones previas. Por ejemplo, el criterio seguido para asignar al hombre el número 1 y a la mujer el número 2 ha sido la tradición metodológica de estudios previos donde se rotuló cada valor de forma caprichosa, tal y como apuntan Morgan y Griego (1998: 25)¹⁵⁰ en un manual de uso sobre SPSS.

Sin embargo, en el caso de la codificación de la variable «Profesión» el valor nominal también indica jerarquía puesto que se ha tenido en cuenta el Real Decreto 1591/2010, de 26 de noviembre, por el que se aprueba la Clasificación Nacional de Ocupaciones 2011. El Instituto Nacional de Estadística realiza tal clasificación para

¹⁵⁰ «Males may be assigned the numeral 1 and females may be coded as 2. This does not imply that females are higher than males or that two males equal a female or any of the other typical mathematical uses of the numeral. The same reasoning applies to many other true nominal categories such as ethnic groups, type of disability, section number in a class schedule, or marital status» (Morgan y Griego, 1998: 25).

asegurar un tratamiento uniforme de los datos estadísticos y realizar comparaciones a nivel nacional e internacional.

3.5.2. Prueba de disponibilidad

La edición de datos del léxico disponible difiere de la encuesta sociológica en tanto que el objetivo no es codificar una variable para interpretarla, sino preservar o alterar las palabras que ha escrito cada informante atendiendo a criterios de edición. Llegados a este punto, cabe plantearse cuáles son los objetivos de las listas de vocabulario que resultarán de la investigación. En el estudio actual podrían resumirse en tres:

- Crear un diccionario con el léxico disponible en español y en lengua extranjera.
- Comparar las listas de palabras resultantes con otras recopiladas en estudios previos.
- Conocer las palabras aportadas por los informantes para realizar análisis estadísticos y cualitativos exhaustivos desde una perspectiva sociolingüística, psicolingüística y de la lingüística aplicada a la enseñanza de idiomas.

Para llevar a cabo la edición de los materiales, se han seguido la mayoría de las pautas y recomendaciones del Proyecto Panhispánico de Léxico Disponible presentes en Samper Padilla (1998) con la incorporación de algunos criterios propios de la investigación. El resultado ha sido una lista completa donde es posible extraer las palabras más disponibles que, a diferencia de los criterios de Samper Padilla (1998) incluye todos los nombres de marcas comerciales y asociaciones indirectas. Esas palabras poco frecuentes o asociaciones más personales permiten conocer de forma más fidedigna el proceso de entrada y salida del vocabulario en el lexicón mental.

Teniendo en cuenta el carácter trilingüe de los datos, se han establecido unos criterios generales de edición que han sido matizados según el idioma y los centros de interés. Esas pautas se subdividen en dos apartados: criterios comunes a la mayoría de estudios herederos del Panhispánico y criterios de edición propios de esta investigación.

3.5.2.1. Criterios comunes al Proyecto Panhispánico de Léxico Disponible

Las siguientes pautas han guiado la edición general de los datos:

- a) Corrección de los errores ortográficos. A pesar de tratarse de alumnado de posgrado se detectaron faltas de ortografía, sobre todo relacionadas con la acentuación en español y palabras en lengua extranjera (inglés o francés). Por ejemplo: *musculo** (*músculo*), *corazon** (*corazón*), *aluvia** (*alubia*), *noise** (*nose*), *profeseur** (*professeur*), entre otras.
- b) Lematización de los sustantivos en la forma neutra. Se han registrado los nombres en masculino singular salvando casos excepcionales en los que el femenino o plural son irregulares o la única forma posible. Es el caso de *matrona*, *chacha*, *gobernanta*, *nun*, *midwife*, *femme de ménage*, etc.
- c) Neutralización de las marcas flexivas. Aparte del género y el número, se han eliminado las marcas verbales (a excepción del gerundio y participio), pasando a la forma de infinitivo. En ciertas expresiones o unidades léxicas se ha conservado el verbo conjugado, por ejemplo *veo veo*.
- d) Uso del paréntesis cuando las unidades léxicas son acortables ya sea porque la unidad léxica esté formada por varias palabras o porque se use la forma abreviada de una sola palabra: *bici(cleta)*, *(auto)bus*, *moto(cyclette)*, *goma (de borrar)*, etc.
- e) Eliminación de las palabras repetidas. Cuando un informante escribe dos veces la misma palabra e incluso hasta tres, se incluye solo la primera vez que aparece la palabra. Este es el caso del informante número dieciséis que escribe: «**brazo**, ojo, pie, uña, dedo, **brazo**, pierna [...]». No obstante, se han conservado los listados originales para realizar otro tipo de análisis desde una perspectiva psicolingüística. Así, resulta imprescindible conocer exactamente qué estímulo concreto evoca una unidad léxica determinada para analizar la relación de las palabras en el lexicón mental.
- f) Unificación ortográfica. En los casos en los que una palabra se puede escribir de manera diferente, se ha optado por la que aparece en los diccionarios

tomados como referencia en cada idioma. Se han recogido palabras escritas de hasta de cuatro formas distintas: *flip-flop*¹⁵¹, *flip-flops*, *flipflop* y *flipflops*.

- g) Lematización de palabras en minúsculas. Como se trata de una lista, es frecuente encontrar todas las palabras con la inicial en mayúsculas ya que el informante entiende que al principio de frase esa ha de ser la regla. Como norma general, solo se ha conservado la inicial mayúscula en el caso de nombres propios, abreviaturas y marcas comerciales.
- h) Eliminación de marcas o signos de puntuación. Con excepción de las tildes, se han eliminado las comillas, el subrayado y los números situados a la izquierda de cada palabra que algunos informantes han añadido a sus listas.
- i) Neutralización de las palabras al grado positivo. Se han perdido los aumentativos y diminutivos no lexicalizados. Sin embargo, en el caso de *barquilla* se ha mantenido ya que se encuentra incluso definida en una entrada distinta en el DRAE.
- j) Inclusión de unidades léxicas formadas por más de una palabra, así como expresiones. Por ejemplo: *salir de marcha* se considera como una única entrada ya que el significado de cada uno de los elementos que componen la estructura es diferente al significado final.

3.5.2.2. Criterios propios

El criterio principal que marca la edición del léxico es la fiabilidad de los datos ya que para adentrarnos en el lexicón mental de cada hablante es necesario conocer las palabras originales que ha utilizado y los errores que ha cometido. Por ello, durante la fase de lematización de las palabras, se han anotado los siguientes apartados para cada informante y centro de interés.

- a) Faltas de ortografía. En español los errores se han clasificado en ocho tipos: 1) acentuación; 2) c, z + e, i; 3) j-g; 4) s-x-c-k-qu; 5) h; 6) b-v; 7) ll-y-ñ; 8) otros. En francés se ha marcado la diferencia entre errores ortográficos debido a la acentuación, ya que existen tres tipos de acentos gráficos diferentes y el resto de disortografías para comparar la proporción con el

¹⁵¹ De esta manera se ha lematizado en este trabajo.

inglés, donde no existen esos acentos. Esta distribución ayuda a valorar si se cumple la hipótesis que la mayoría de las faltas de ortografía se deben a errores de acentuación.

- b) Plurales y femeninos. Se ha recogido el número de palabras que el informante ha escrito en plural. De esta forma, se han localizado palabras siempre escritas en plural, tales como, *extremidades*. En el caso del femenino, se ha realizado un recuento por separado en el centro 8, denominado *Profesiones y oficios*. Ha sido común encontrar profesiones más frecuentes en femenino, como *enfermera*, o solo existentes en femenino, como *matrona* o *chacha*. Las palabras heterónimas se han mantenido en sus dos formas, tal y como propone Samper Padilla (1998: 315). Por ende, los plurales irregulares, presentes sobre todo en lengua extranjera, como *œil-yeux* en francés, han sido recogidos de forma diferenciada.
- c) Coloquialismos. Se han añadido palabras de la lengua común que a menudo no aparecen en los diccionarios o si aparecen no cuentan con esa acepción. Este es el caso de *nube* cuando se refiere a *café* ya que se encuentra registrado en el centro *Comidas y bebidas*.
- d) Palabras extranjeras. Las palabras procedentes de otros idiomas, estén acuñadas o no, han sido escritas bajo la forma aconsejada en español o en su forma original, consultando previamente si existe alguna aclaración al respecto propuesta por la Real Academia Española. Cuando se refieren a realidades similares, como *parking* y *aparcamiento*, se han mantenido las dos formas. Además de contener connotaciones diferentes, *parking* suele referirse a un aparcamiento privado y a menudo de pago mientras que *aparcamiento* es una palabra mucho más genérica. Por ello, si se cambia y lematiza todo al español los resultados no reflejarían el léxico realmente utilizado por los informantes. Para la clasificación de los extranjerismos, se han tenido en cuenta las recomendaciones de la *Ortografía de la Lengua Española* (2010) y del *Diccionario Panhispánico de Dudas*¹⁵². En este diccionario se aplicaron los siguientes criterios con el fin de encontrar soluciones gráficas o fonéticas a la incorporación de

¹⁵² Versión electrónica: <http://www.rae.es/recursos/diccionarios/dpd> [Fecha de última consulta el 20 de marzo de 2015].

aportaciones léxicas de otras lenguas en español distinguiendo entre: extranjerismos superfluos o innecesarios, para los que se proponen palabras equivalentes en español; extranjerismos necesarios o muy extendidos en los que se opta, bien por mantener la grafía y pronunciación originarias escritos en cursiva, o bien por adaptar la pronunciación o la grafía al español y se escriben en letra redonda.

- e) Nombres de marcas comerciales y sustantivos propios. A diferencia de la recomendación de Samper Padilla (1998: 317) sobre la eliminación de los lexemas que no estuviesen lexicalizados e incorporados al lenguaje común, en esta investigación se han incluido todos los nombres de marcas comerciales y otros nombres propios. Este criterio es el seguido por Mateo García (1998: 86), Galloso Camacho (2003: 30), Gómez Molina y Gómez Devís (2004: 52), Ávila Muñoz (2006: 56), Ávila Muñoz y Villena Ponsoda (2010: 94), entre otros. Además, los estudios de Málaga (Ávila Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010) incorporan una premisa: la marca debe ser escrita por, al menos, 10 informantes. En este estudio se ha considerado que los lexemas que no forman parte del vocabulario común, fruto de asociaciones particulares, tendrán un bajo nivel de disponibilidad, por lo que su inclusión supone un enriquecimiento y no un impedimento para construir las listas del léxico disponible. De esta forma, es posible conocer realmente cuáles están lexicalizados, tal y como expone Galloso Camacho (2003: 30).
- f) Asociaciones individuales. Resulta muy complejo saber si una determinada palabra es una asociación individual si no se recoge y analiza posteriormente. Al igual que en Galloso Camacho (2002a), Ávila Muñoz (2006) y Ávila Muñoz y Villena Ponsoda (2010) se han recogido todos los vocablos aportados por los informantes, de forma que se pudiesen realizar diferentes tipos de análisis sin caer en la manipulación de datos como advierten Mateo García (1998: 86), Galloso Camacho (2002a: 27) y Ávila Muñoz (2006: 58).
- g) Uso de la coma. Algunos nombres de juegos como es el caso de *pie*dra, *pa*pel o *tij*era requieren el uso de la coma. Sin embargo, al añadir la coma el programa Lexidisp lematiza como palabras diferentes *pie*dra y, por otro lado, *pa*pel o *tij*era.

Cabe recordar que la marca utilizada para diferenciar una palabra de otra en todos los casos es una coma, de ahí que presentase ese error de edición. Para subsanarlo, se optó por escribir un punto para separar las palabras, *piedra. papel o tijera*, y una vez incorporado el juego a la lista se sustituyeron todos los puntos por comas, obteniendo la forma correcta en el diccionario de léxico disponible.

- h) Notas. En este punto se han recogido aclaraciones y anotaciones que pueden ser susceptibles de análisis. Los apuntes más relevantes serán comentados con mayor detenimiento en los análisis, como es el caso de las faltas de ortografía¹⁵³.

3.5.2.3. Características de la edición en lengua española

Las palabras han sido consultadas en el *Diccionario de la lengua española*, concretamente la 22.^a edición publicada en 2001¹⁵⁴. Uno de los apartados recogidos en cada centro de interés han sido los extranjerismos estableciendo la diferencia entre los que se adaptan a las convenciones del español y aparecen en letra redonda (incluyendo los derivados españoles) y aquellos cuya representación gráfica o pronunciación no se ajustan a los usos del español. Además, se han tenido en cuenta los cambios propuestos para la vigésima tercera edición¹⁵⁵. A continuación, se comentan los rasgos principales de la edición de cada centro de interés:

Centro 1. *Partes del cuerpo humano*

- Uso del paréntesis en vocablos que a veces aparecen escritos en su forma abreviada o con una de las dos palabras que lo forman: *columna (vertebral)*, *(dedo) índice*, *(dedo) pulgar*, *palma (de la mano)*, *planta (del pie)*. Sin embargo, ningún informante ha escrito *dedo meñique*, por lo que se ha lematizado como *meñique*. Lo mismo ocurre con otras palabras como *esternocleidomastoideo*, *yugular* y *aorta* pues no hay ningún registro precedido de músculo ni vena.

¹⁵³ Vid. apartado 4.6.

¹⁵⁴ La fecha de última consulta en línea de las diferentes definiciones dadas en este capítulo se han realizado el 20 de marzo de 2015, de ahí que no vengán especificadas en cada caso para evitar su redundancia.

¹⁵⁵ Disponible en línea en: <http://www.rae.es/recursos/diccionarios/drae> [Fecha de última consulta el 20 de marzo de 2015].

- La palabra *isquiotibial* no aparece en el DRAE. Sin embargo, en diferentes artículos médicos aparece que el *isquiotibial* está compuesto por un conjunto de tres músculos, de acuerdo con su denominación, comienza en el *isquion* y termina en la *tibia*.
- Palabra en plural: en el DRAE *caries* solo aparecen en plural.
- *Canino*: es un adjetivo que se refiere a diente. En el DRAE aparece como un compuesto, razón por la cual se ha lematizado como *diente canino*.

Centro 2. *La ropa*

- Palabras de origen extranjero aceptadas por el DRAE: *anorak*, *corsé*, *chaqué*, *chilaba*, *culote*, *frac*, *fular* y *jersey*.
- Extranjerismos que el DRAE acepta en cursiva y que, por tanto, mantienen la grafía y pronunciación originarias: *blazer*, *body*, *panty*, *slip* y *short* (está propuesto para la supresión en la vigésima tercera edición del diccionario).
- Palabras de origen extranjero no aceptadas por el DRAE que, por lo tanto, están en cursiva en el diccionario: *bóxer*, *burka*, *burkini*, *jeans*, *leggings*, *pashmina* y *triquini*.
- Marcas: *Borsalino* y *Wonderbra* son marcas que se utilizan para hablar de un tipo concreto de sombrero y de sujetador, pero al no estar ese uso acuñado, se han escrito en mayúsculas.

Centro 3. *Comidas y bebidas*

- Marcas comerciales: este centro de interés es el segundo en número de marcas aportadas por los informantes, tras el noveno, *Ordenadores e internet*. Se han incluido las siguientes: *Aquarius*, *Bacardi*, *Cacique*, *Chardonnay* (aunque no es realmente una marca, sino un tipo de uva), *Coca-Cola*, *Cola Cao*, *Danone*, *Danup*, *Doritos*, *Dyc*, *Fanta*, *Fanta Naranja*, *Fanta Limón*, *Gominolas*, *JB*, *Seven Up*. Cabe destacar que *Gominolas* se ha lematizado como marca ya que el DRAE no lo recoge como sustantivo común a pesar de que sea una palabra frecuentemente usada.
- Palabras de origen extranjero aceptadas por el DRAE: *beicon*, *caqui*, *canapé*, *chupachús* (propuesta para la próxima edición), *cuscús*, *coñac*, *entrecot*, *cruasán*, *fidená*, *kebab*, *mojito*, *bistec*, *roquefort*, *sándwich*, *yogur*, *sake* y *vodka*.

- Palabras que han sido aceptadas por el DRAE pero en cursiva: *baguette*, *brandy*, *carpaccio*, *curry*, *foie gras*, *gin-tonic*, *ketchup*, *mozzarella* y *pizza*.
- Palabras que no han sido aceptadas: *brioche*, *calzone*, *tomate cherry*, *crêpe* (el diccionario recoge *crepé* pero referido a un tejido), *fusilli*, *limoncello*, *muffin*, *tortellini* y *wrap*.
- Localismos: *calamar* y *calamarito* en Málaga se refieren a realidades diferentes ya que el tamaño del *calamarito* como su nombre indica es menor¹⁵⁶. *Calamar* no se encuentra en el listado de léxico disponible. Esta palabra ha sido buscada en la base de datos ICIOTERM¹⁵⁷ donde se recogen la denominaciones locales del término científico *Allotienthis subulata*¹⁵⁸.

En el caso de *tintorera* es diferente ya que el DRAE lo recoge haciendo referencia a una especie de tiburón presente en las costas del sur de España y Marruecos, pero no se designa a la especie que realmente se cocina en Andalucía. El diccionario científico lo reconoce bajo el nombre *prionace glauca*¹⁵⁹ y su tamaño puede llegar a cinco centímetros de longitud, muy diferente a los tres o cuatro metros que según el DRAE puede llegar a alcanzar. Por tanto, se trataría también de un localismo ya que la definición del DRAE no se corresponde a la realidad del pez en Málaga.

Nube también se utiliza solo en Málaga para designar un café corto con bastante leche. El *Tesoro Léxico de las Hablas Andaluzas*, en su novena acepción lo define como «Café con leche en el que apenas hay café. [VPM: Málaga]» (Alvar Ezquerro, 2000: 555). La descripción más pormenorizada de este localismo se encuentra en la publicidad de la conocida marca de cafés en Málaga, Santa Cristina, la cual establece que la nube debe servirse con 10 % de

¹⁵⁶ «Calamarito (m.) Cría del calamar. [LMP, 686: Almuñécar (Gr.), Motril (Gr.); Estepona (Má.), Málaga]» (Alvar Ezquerro, 2000: 176).

¹⁵⁷ ICIOTERM es una base de datos de ictieterminológicos asociados a una correcta y actual identificación de especies pesqueras de las costas de Andalucía, resultante de dos Proyectos de Investigación llevados a cabo por investigadores del Instituto de Ciencias Marinas de Andalucía del Consejo Superior de Investigaciones Científicas y del Departamento de Filología y Traducción de la Universidad Pablo de Olavide de Sevilla. La descripción del proyecto se encuentra en: http://www.ictieterm.es/proyecto_introduccion.php [Fecha de última consulta el 20 de febrero de 2015].

¹⁵⁸ Para más información sobre el término, véase el siguiente enlace: http://www.ictieterm.es/nombre_cientifico.php?nc=22 [Fecha de última consulta el 20 de febrero de 2015].

¹⁵⁹ Ulteriores detalles disponibles en: http://www.ictieterm.es/nombre_cientifico.php?nc=10 [Fecha de última consulta el 20 de febrero de 2015].

café y un 90 % de leche¹⁶⁰. También aparece una entrada en el *Diccionario del habla malagueña* documentado por Enrique del Pino (2006) que lo define como «café con leche en el que la proporción de café es mínima».

Espeto redirige en el DRAE a *espetón*. En su sexta acepción, aparece marcado con un significado diferente en Andalucía y lo define como sigue: «*And.* Conjunto de sardinas que se atraviesan con una caña para asarlas».

- Palabras en plural: *callos* y *migas* se escriben en plural cuando se refieren a comida.
- Nombres propios: *Brie* se escribe con mayúsculas porque designa el lugar de procedencia del queso; *Sex on the Beach* se refiere a un tipo de cóctel; *Ribera del Duero* y *Sauvignon* son las denominaciones de un vino.

Centro 4. *La escuela: muebles y materiales*

- Uso del paréntesis en los siguientes vocablos: *agenda (escolar)*, *autobús (escolar)*, *boli(grafo)*, *goma (de borrar)*, *lápiz (de cera)*, *(ordenador) portátil*.
- *Aprobado* y *aprobar* aparecen como dos formas diferentes ya que *aprobado* puede hacer referencia a una nota específica (de 5 a 6 puntos) y *aprobar* se refiere al verbo en general independientemente de la nota conseguida.
- Los nombres de asignaturas se han lematizado en minúsculas ya que puede referirse a la ciencia en general o al nombre de la asignatura en concreto: *asignatura*, *matemáticas*, *inglés*, *francés*, *lengua castellana y literatura*, *biología*, *geología*, *latín*, *griego*, *filosofía*, *química*, *física*, *educación física*, *educación plástica*, *música*.
- Marcas: *Fixo*, *Tipp-Ex*, *Pilot* y *Post-it*.
- Neologismos en cursiva no recogidos en el DRAE: *pendrive* y *netbook*.
- Palabras muy usadas que no aparecen en el DRAE: *quita-grapas*, *subrayador* y *porta-ángulos* (el diccionario recoge *porta-* como elemento para formar palabras compuestas).

Centro 5. *La ciudad*

- Nombres propios: *Alliance Française*, *AVE*, *Burger King*, *Cercanías*, *Emasa*, *El Retiro*, *Endesa*, *EMT* y *Renfe*.

¹⁶⁰ En <http://www.productosdelcafe.com/cafessantacristina/cafemalaga.html#> [Fecha de última consulta el 20 de febrero de 2015].

- Topónimos: *Barcelona, Madrid, Málaga y Nueva York*. Por analogía con los nombres propios, se mantienen los nombres de ciudades. Además, algunas editoriales recogen los nombres de ciudades más famosas en los léxicos básicos puesto que algunas suelen cambiar la grafía según el idioma.
- Palabras en plural: el DRAE define *afueras* como los alrededores de una población, por lo que se ha lematizado en su forma plural. El vocablo *decomisos* utiliza la forma de plural como si se tratase de singular. El DRAE en su cuarta acepción¹⁶¹ recoge esta palabra en forma plural, aportando el siguiente ejemplo «El decomisos de la esquina abre los sábados».
- *Estación de autobús, de metro y de tren* aparece con el medio de transporte en singular porque el DRAE describe en la entrada de estación lo siguiente: edificio o edificios en que están las oficinas y dependencias de una estación del ferrocarril o de autobús.
- *La Hiperronda de Málaga* (MA-40) es el nombre que se le da a la Segunda Circunvalación de Málaga.
- Palabras extranjeras admitidas por el DRAE: *bulevar, chalé, club, videoclub* y *vip*.
- Palabras extranjeras admitidas por el DRAE en cursiva: *boutique* y *souvenir*.
- Palabras extranjeras no admitidas por el DRAE: *cibercafé, outlet, pub* y *parking* (se ha incluido porque en español se utiliza de forma diferenciada a *aparcamiento*, por la connotación comentada previamente que tiene *parking* como *aparcamiento* privado y normalmente de pago).
- Locuciones: se ha lematizado *ceda el paso* como si fuese un sustantivo ya que se refiere a la señal de tráfico.

Centro 6. *Medios de transporte*

- Asociaciones individuales: el *AT-AT* (All Terrain Armored Transport) hace referencia a uno de los vehículos imperiales de enorme tamaño con característicos rasgos zoomorfos, propios de la saga *Star Wars* y el *AT-ST* (All Terrain Scout Transport) es una máquina de combate bípeda gigante controlada por humanos del mundo ficticio de *Star Wars*.

¹⁶¹ «4. m. pl. u. c. sing. coloq. Establecimiento autorizado donde se venden mercancías decomisadas».

- Palabras extranjeras que el DRAE admite: *kayak*.
- Palabras extranjeras que el DRAE admite en cursiva: *ferry*, *jet* y *scooter*.
- Palabras extranjeras que el DRAE no admite: *parking*, *quad*, *tuc tuc* y *vaporetto*.
- Marcas: *JCB* (marca de excavadoras), *Mini* y *Seiscientos*.
- Nombres propios: *AVE*, *Cercanías* y *Renfe*.
- Se han lematizado en entradas diferentes *moto acuática* y *moto de agua* ya que ambos términos se utilizan de manera autónoma.
- Palabras recogidas como avances de la vigésimo tercera edición del DRAE consultada en línea: *zodiac* y *vespa*.

Centro 7. Juegos y distracciones

- Palabras extranjeras admitidas por el DRAE: *aerobic*, *bádminton*, *test*, *kárate*, *karaoke*, *kayak*, *puzzle*, *surf* (sin embargo, no aparece el verbo *surfear*), *taichi*¹⁶², *taekwondo*, *triatlón* y *waterpolo*.
- Palabras extranjeras admitidas por el DRAE en cursiva: *camping*, *cricket*, *footing*, *hockey*, *jogging*, *paddle*, *ping-pong* y *rugby*.
- Palabras extranjeras no admitidas por el DRAE: *aquagym*, *balconing* (aunque tampoco está admitida en los diccionarios de lengua inglesa), *blog*, *chat*, *chatear* (existe con otro significado), *fitness*, *freaky*, *funky*, *paintball*, *pub*, *puenting*, *rafting*, *snowboard*, *spinning*, *squash*, *step*, *sudoku* y *trekking*.
- Nombres propios de juegos: *Bubble Bobble*, *Buscaminas*, *Buzz!*, *¿Dónde está Wally?*, *Dicciopinta*, *El Gordo*, *Farm Ville*, *Game Boy*, *Jungle Speed*, *Mario Bros*, *Monopoly*, *Party & Co.*, *Pet Society*, *Play Mobil*, *Play Station*, *Pokemon*, *¿Quién es Quién?*, *Scattergories*, *Sega Mega Drive*, *Sonic*, *Super Pang*, *Tabú*, *Tetris*, *Trivial*, *UNO*, *Wii*, *World of Warcraft*, *Yathzee* y *Zelda*¹⁶³.
- Páginas web: *Google*, *Hotmail*, *Messenger*, *Pro Evolution Soccer 2011*, *Risk*, *Series Yonkis* y *Wikipedia*.

¹⁶²Al elegir en el buscador de la Real Academia Española entre los diferentes diccionarios, en el *Diccionario de la lengua española* aparece *taichi* sin tilde a pesar de ser una palabra aguda. Sin embargo, en el *Diccionario esencial* y en el *Diccionario panhispánico de dudas* la palabra *taichí* está acentuada.

¹⁶³ Este nombre viene del juego de la Nintendo llamado *The Legend of Zelda*. Sin embargo, en español se suele traducir (aunque la marca no lo haga) o bien acortarlo, *Zelda*.

- Se ha lematizado *Buzz!* con signo de exclamación al final ya que así está registrada la marca, de no poner la exclamación podría confundirse con *Buzz*, referido al canal de televisión, y *Google Buzz*, el mecanismo de red social de Google.
- Nombres de juegos recogidos en el DRAE en plural: *bolos*¹⁶⁴, *canicas*¹⁶⁵, *chapas*¹⁶⁶, *casitas*, *damas*¹⁶⁷, *palmas* y *pesas*¹⁶⁸. Cabe mencionar que todos los vocablos suelen ir precedidos de *jugar* excepto *pesas* que lleva asociado el verbo *hacer*.
- Se ha lematizado de forma diferente *UNO* y *uno* ya que hacen referencia a dos juegos diferentes. El primero es un juego popular de cartas producto de Mattel desde 1992 pero desarrollado en 1971 por Merle Robbins en Reading, Ohio. El segundo es un juego de naipes al que se juega tradicionalmente con una baraja española completa o con una baraja inglesa.

Centro 8. Profesiones y oficios

- Palabras extranjeras admitidas por el DRAE: *barman*, *bróker*, *chef* y *chófer*.
- Palabras extranjeras admitidas por el DRAE en cursiva: *au pair*, *disc-jockey*, *maître* y *top-model*.
- Palabras no admitas por el DRAE: *stripper*¹⁶⁹, *hacker*. Cabe mencionar que el *Diccionario panhispánico* redirecciona *hacker* a *pirata informático* pero en este caso se han lematizado como dos palabras diferentes ya que de esta forma es posible conocer si es más usual la voz inglesa o española.

¹⁶⁴ Vigésima acepción de la palabra. «m. pl. Juego que consiste en poner derechos sobre el suelo cierto número de bolos y derribar cada jugador los que pueda, arrojándoles sucesivamente las bolas que correspondan por jugada».

¹⁶⁵ Primera acepción de la palabra. «Juego de niños que se hace con bolas pequeñas de barro, vidrio u otra materia dura. U. m. en pl».

¹⁶⁶ Decimoséptima acepción. «f. pl. Juego entre dos o más personas, que consiste en tirar por alto dos monedas iguales. Si al caer al suelo quedan ambas con la cara hacia arriba, el que las ha tirado gana a todos y sigue tirando; en caso contrario paga todas las puestas y deja de tirar; y si resulta cara y cruz, ni pierde ni gana, y tira de nuevo». Decimotava acepción. «f. pl. Juego infantil en que se utilizan las chapas de las botellas».

¹⁶⁷ Novena acepción. «f. pl. Juego que se ejecuta en un tablero de 64 escaques, con dos conjuntos de fichas distinguidos por el color».

¹⁶⁸ Tercera acepción. «f. Dep. Pieza muy pesada que se emplea en halterofilia o para hacer gimnasia. U. m. en pl.

¹⁶⁹ El *Diccionario panhispánico* recomienda adaptar la palabra al español, teniendo como resultado *estriper*, pero los informantes lo han escrito en inglés».

- Palabras con forma plural usadas en singular y plural: *botones*¹⁷⁰ y *relaciones públicas*¹⁷¹.
- Palabras en femenino: *chacha*, definido en el DRAE como sirvienta; *matrona*, ya que no existe la forma masculina; *niñera*¹⁷² por contar con diferentes significados en femenino y masculino; *costurera*, ya que el DRAE ha escrito que en masculino está anticuado, proponiendo como variante *modista* y *sastre*; *gobernanta*¹⁷³, en el DRAE está recogido en femenino con diferente significado a *gobernante*¹⁷⁴.
- *Yesista* y *yesero* son dos palabras comunes en cuanto a significado pero de diferente derivación, por lo que se han mantenido ambas formas.
- *Papa* ha sido lematizado en minúsculas ya que el DRAE no especifica nada y lo define escrito en minúsculas como: «1. m. Sumo Pontífice romano, vicario de Cristo, sucesor de San Pedro en el gobierno universal de la Iglesia católica, de la cual es cabeza visible, y padre espiritual de todos los fieles».
- Se ha diferenciado la denominación del docente en las etapas educativas: profesor de primaria, profesor de secundaria y profesor universitario.
- Se han dejado tal cual las expresiones: *salir de marcha* y *salir de fiesta*.
- *Teleoperador* no viene recogido en el DRAE pero el primer lexema aparece como un elemento de composición que significa «a distancia», como puede ser el caso de teléfono, televisión, telediario, telespectador, entre otras.

Centro 9. Ordenadores e internet¹⁷⁵

- Palabras extranjeras admitidas por el DRAE: *bit*, *web*, *campus*, *tóner*, *casete* y *chip*.
- Palabras extranjeras admitidas por el DRAE en cursiva: *byte*, *hardware* y *software*.

¹⁷⁰ Decimosexta acepción. «(Por los botones que suele lucir en su uniforme). m. pl. m. pl. Muchacho que sirve en hoteles y otros establecimientos para llevar los recados u otras comisiones que se le encargan».

¹⁷¹ Segunda acepción para relaciones públicas encontrado bajo el lema relación. «f. pl. u. c. sing. com. Persona que desempeña esta profesión».

¹⁷² Segunda acepción. «f. Criada destinada a cuidar niños. En masculino el DRAE solo incluye la siguiente acepción: 1. adj. Que gusta de niños o de niñerías».

¹⁷³ Primera acepción. «f. Mujer que en los grandes hoteles tiene a su cargo el servicio de un piso en lo tocante a limpieza de habitaciones, conservación del mobiliario, alfombras y demás enseres». Segunda acepción. «f. Encargada de la administración de una casa o institución».

¹⁷⁴ Primera acepción. «adj. Que gobierna. Partido gobernante. U. m. c. s.». Segunda acepción. «m. coloq. Hombre que se mete a gobernar algo».

¹⁷⁵ Este centro se caracteriza por la gran cantidad de neologismos y nombres propios, por lo que además de incluir las palabras buscadas en el DRAE se ha consultado el *Diccionario panhispánico*.

- Palabras extranjeras admitidas en el Panhispánico pero no en el DRAE: *e-mail*, *chatear* y *hacker* (pero no aparece *hackear*).
- Palabras extranjeras no admitidas ni por el Panhispánico ni por el DRAE: *bluetooth*, *blog*, *cookie*, *firewall*, *freaky*, *router*, *hackear*, *hipervínculo* y *pendrive*.
- Nombres propios: *Access*, *Acer*, *Android*, *Ares*, *Audacity*, *AutoCAD*, *Avast*, *Dell*, *DNS*, *E-Mule*, *Intel Core*, *FireWire*, *Firefox* y *Java*.
- *Ajax*: acrónimo de Asynchronous JavaScript And XML (por lo tanto se escribe la primera en mayúscula).
- La palabra *antiespía* no la recoge el DRAE. Por similitud a otras palabras con el prefijo *anti-* se ha escrito junta, como en el caso de *antivirus*.
- Las extensiones de las imágenes más habitualmente utilizadas como *BIN* y *JPG* se han añadido.
- C++ es un lenguaje de programación diseñado a mediados de los años 1980 por Bjarne Stroustrup. La expresión C++ significa «incremento de C» y se refiere a que C++ es una extensión de C.
- *Ciberacoso* se ha lematizado como una sola forma ya que en el DRAE se recoge el lexema *ciber-* como elemento de composición presente en palabras como *cibernáutico* y *ciberespacio*.
- *FriendFeed*: aparecen dos mayúsculas diferentes en una misma palabra para mantener así la forma original del nombre propio.
- *Gateway* puede denominar el nombre de una empresa o a la traducción en español de pasarela o puerta de enlace. En este caso, se refiere a un dispositivo, con frecuencia una computadora, que permite interconectar redes con protocolos y arquitecturas diferentes.
- *HDMI* hace referencia a una interfaz multimedia de alta definición y se utiliza comúnmente en español.

3.5.2.4. Características de la edición en lengua extranjera

La lematización en inglés y francés toma como referencia investigaciones previas de disponibilidad léxica en español como lengua extranjera y como segunda

lengua: Carcedo González (1998a, 2000c), Samper Hernández (2002), Frey Pereira (2006), Sánchez-Sauz Laserna (2011), Gallego Gallego (2011, 2014) y Fernández-Merino Gutiérrez (2015), entre otras. Asimismo, se tienen en cuenta investigaciones previas en inglés como lengua extranjera, tales como Germany y Cartes (2000) y Ferreira (2006).

La peculiaridad de la edición de datos en lengua extranjera reside en las faltas de ortografía, ya que en ocasiones hay que intuir a qué palabra se refiere el informante. Al igual que en Samper Hernández (2002: 28), se ha tenido en cuenta el campo léxico al que pertenecen las palabras, los vocablos que le preceden y proceden y la posición que ocupan en la lista. Sin embargo, en este estudio la lengua materna de los informantes es el español por lo que es más fácil deducir la palabra correcta. Según el caso, el error se ha debido a factores diversos destacando las interferencias con el español y la confusión entre la forma de pronunciar un vocablo y escribirlo. Por ejemplo, el término *nose* se pronuncia en inglés /nəʊz/ por lo que un error típico es escribirlo con un diptongo, *noise* cuyo significado es ruido. Samper Hernández (2002: 23) afirma que «[...] al tratarse informantes no nativos, nos encontramos con fenómenos tales como la interferencia de la lengua materna de los hablantes o la semejanza de una unidad determinada a una palabra española sin llegar a pertenecer sin embargo a nuestro idioma».

En lengua inglesa, el diccionario bilingüe utilizado para la traducción de las palabras es el *Collins Spanish Dictionary*¹⁷⁶ y el diccionario monolingüe de referencia es el *Collins Concise English Dictionary*¹⁷⁷. La búsqueda de palabras en francés se ha llevado a cabo en el diccionario monolingüe *Trésor de la Langue Française* de consulta libre en internet¹⁷⁸. En cuanto al diccionario bilingüe, se ha usado la versión en línea del *Diccionario Espasa Grand: español-francés français-espagnol* de la editorial Espasa-Calpe¹⁷⁹.

a) Edición en lengua inglesa

Centro 1. *Human body parts*

¹⁷⁶ *Collins Spanish Dictionary. Complete and Unabridged.* (2005). (8th Edition). Glasgow: HarperCollins Publishers Ltd.

¹⁷⁷ *Collins Concise English Dictionary.* (2008). (7th Edition). Glasgow: HarperCollins Publishers Ltd.

¹⁷⁸ Véase: <http://atilf.atilf.fr>. [Fecha de última consulta el 20 de febrero de 2015].

¹⁷⁹ Disponible en: <http://www.wordreference.com/fres> [Fecha de última consulta el 20 de febrero de 2015].

- Se han incorporado los plurales irregulares: *foot-feet* y *tooth-teeth*.
- Errores típicos: *mouse* o *mauth* en lugar de *mouth*, *noise* en vez de *nose*; *knew* o *kenie* para decir *knee*; *lib* en lugar de *lip*, *chess* (ajedrez) para decir *chest*; *eyebrown* en vez de *eyebrow*.
- Palabras entre paréntesis: *(eye)lash*.
- Palabras clasificadas como vulgares en el diccionario: *ball*, *belly*, *belly button*, *chick*, *cock*, *cunt*, *dick*, *pussy*, *tit* y *tummy*.
- Se han lematizado de forma diferente palabras procedentes del inglés británico como es el caso de *botton* y del inglés americano *butt*.

Centro 2. *Clothes*

- Palabras lematizadas al plural: *accesories*¹⁸⁰, *boxer shorts*¹⁸¹, *briefs*¹⁸², *chinos*¹⁸³, *dungarees*¹⁸⁴, *French sneakers*, *glasses*¹⁸⁵, *high heels*, *hot pants*¹⁸⁶, *jeans*, *knickers*¹⁸⁷, *leggings*¹⁸⁸, *panties*¹⁸⁹, *knickers*, *pants*¹⁹⁰, *pyjamas*¹⁹¹, *rompers*¹⁹², *shorts*¹⁹³, *sunglasses* y *tights*¹⁹⁴.
- Falsos amigos: las palabras *body* y *body stocking* se han lematizado de forma diferente. La primera significa cuerpo y la segunda se refiere a bodi en español, aunque presumiblemente el informante que ha escrito *body* haya querido decir en lengua materna *bodi* ya que aparece en el centro referido a la ropa.
- Palabras lematizadas de forma diferente que hacen referencia a una misma realidad: *high-heeled shoe* y *high heels*.

¹⁸⁰ «Plural noun: a small article carried or worn to complement a garment».

¹⁸¹ «Plural noun: men's underpants resembling shorts».

¹⁸² «Plural noun: short, close-fitting underpants or knickers».

¹⁸³ «Pl.: casual trousers made from chino».

¹⁸⁴ «Brit.: trousers with a bib held up by straps over the shoulders, made of denim or a similar material; Amer. hard-wearing denim trousers».

¹⁸⁵ «Plural noun: a pair of lenses set in a frame that rests on the nose and ears, used to correct or assist defective eyesight».

¹⁸⁶ «Plural noun: women's very tight, brief shorts».

¹⁸⁷ «Plural noun Brit. a woman's or girl's undergarment, covering the body from the waist or hips to the top of the thighs and having two holes for the legs».

¹⁸⁸ «Plural noun: 1 tight-fitting stretchy trousers worn by women».

¹⁸⁹ «Plural noun: informal legless underpants worn by women and girls».

¹⁹⁰ «Plural noun: Brit. underpants or knickers; Amer. trousers».

¹⁹¹ «Plural noun: a jacket and loose trousers for sleeping in».

¹⁹² «Plural noun: a young child's one-piece outer garment».

¹⁹³ «Plural noun: short trousers that reach only to the knees or thighs; Amer. men's underpants».

¹⁹⁴ «Plural noun a woman's close-fitting garment made of a knitted yarn, covering the legs, hips, and bottom».

- Unificación de términos bajo la misma forma: *flip flop*, *flip flops*, *flip-flop* y *flip-flops* bajo el lema *flip-flop*¹⁹⁵; *minishort* y *mini short* escritos como *mini short*; *mini skirt* y *miniskirt* lematizados como *miniskirt*; *pull over* y *pullover* lematizados como *pullover*.

Centro 3. *Food and drink*

- Palabras británicas: *alcopop* definida como «noun Brit. informal a ready-mixed soft drink containing alcohol».
- Errores frecuentes: *bear* (oso) en lugar de *beer* (cerveza); *McDonald's* escrito como *MacDonals* o *Mac Donald's*; *pop corn* en lugar de *popcorn*; *rib eye* se aconseja escribir con un guión, *rib-eye*; *well done* es un calco del español, mejor escribir *well-cooked*; *yougurt* en vez de *yogurt*; *tomatoe* en lugar de *tomato*; *peper* en el caso de *pepper* y *oignon* para decir *onion*.
- Unificación de grafías: *chamomile* y *camomile* bajo el lema *chamomile*.
- Palabras en plural: *cornflakes*¹⁹⁶ y *French fries*.
- Palabras de origen extranjero: *crème caramel*, *kir*, *panacotta*, *piz̧za*, *sangria*, *spaghetti* y *tortellini*.

Centro 4. *The school: furniture and school material*

- Falsos amigos: *agenda* se traduce como orden del día; seguramente, el informante habrá querido decir *diary* cuya traducción en español es agenda.
- Errores frecuentes: *encyclopedia* en lugar de la forma correcta *encyclopaedia*; *gymnastics* y *physics* tienden a escribirse sin la -s final, como en español. Sin embargo, cuando se refiere a asignatura tiene que llevar la marca de plural aunque reciba el mismo trato que en singular;¹⁹⁷ *sisors* en vez de *scissors*; *lap top* en el caso de *laptop*; *market* (mercado) en lugar de *marker* (rotulador); *subjet* para decir *subject*; *text book* en lugar de *textbook*.
- Errores en marcas: *post it* se edita *Post-it*; *play dough* en lugar de *Play-Doh* y *selotape* en lugar de *Sellotape*.

¹⁹⁵ «Noun: a light sandal with a thong that passes between the big and second toes».

¹⁹⁶ «Plural noun: a breakfast cereal consisting of toasted flakes made from maize flour».

¹⁹⁷ «Plural noun [treated as sing.] the branch of science concerned with the nature and properties of matter and energy».

Centro 5. *The city*

- Palabras británicas: *car park*¹⁹⁸, *litter bin* y *semi-detached*¹⁹⁹.
- Palabras en plural: *outskirts*²⁰⁰ se escribe en plural al igual que ocurría con su equivalente en español, *afueras*.
- Gran cantidad de compuestos con el lexema *shop*: *barber's shop*, *butcher's shop*, *chemist's shop*, *sex shop*, *shopping centre* y *shopping mall*.
- Nombres de ciudades: *Benalmádena*, *Málaga*, *New York*, *London* y *Torremolinos*.

Centro 6. *Means of transport*

- Este centro suele ser el menos productivo y además cuenta con gran cantidad de vocablos relacionados de forma indirecta, tales como: *by foot*, *by walk*, *water*, *bus station*, *bus stop*, *go home*, *go to job*, *footing*, *jogging*, *sky*, *ticket*, *shoe*, *oil*, *oil station*, etc.

Centro 7. *Games and entertainments*

- Palabras en plural: *aerobic* en español se escribe en singular pero en inglés cuando hace referencia al sustantivo se escribe en plural, *aerobics*.
- Nombres propios: *Dance Revolution*, *Dogz*, *Facebook*, *Friends*, *How I met your mother*, *Just Dance*, *Lost*, *Mario Bros*, *Monkey Island*, *Monopoly*, *Moto Racer*, *Nintendo*, *Playmobile*, *Play Station*, *Rabbidz*, *Revolution*, *Scrabble*, *Sims*, *Sky*, *Skype*, *Starbucs*, *Taboo*, *The Legend of Kyrandia*, *Trivial*, *Tuenti*, *Twitter*, *Wii*, *World of Warcraft* y *Xbox*.

Centro 8. *Professions and jobs*

- Palabras escritas en la forma femenina: *actress*, *businesswoman*, *firewoman*, *headmistress*, *housewife*, *policewoman*, *midwife*, *nun*, *policewoman*, *postwoman*, *sportswoman*, *stewardess* y *waitress*.
- Unificación de palabras: *baby-sitter* y *babysitter* se han lematizado como *babysitter*; al igual que *house-husband* y *househusband* como *househusband*.

¹⁹⁸ «Brit.: an area or building where cars or other vehicles may be left temporarily».

¹⁹⁹ «Brit.: (of a house) joined to another house on one side only by a common wall. House».

²⁰⁰ «Plural noun: the outer parts of a town or city».

Centro 9. *Computer and Internet*²⁰¹

- Gran cantidad de nombres propios: *Access, Apple, Bill Gates, Dropbox, eMule, Excel, Exe-Learning, Explorer, Facebook, Firefox, Gran hermano, Google, Google Chrome, Google Maps, Guadalinex, Herald Tribune, Hotmail, Hot Potatoes, Java, Mackintosh, Messenger, Microsoft, MSN, Photoshop, The times, Tuenti, Twitter, Seneca, Skype, Slideshare, SPSS, SUDDAAM, My Space, Linked in, Series Yonkies, Spotify, Wikipedia, Windows, Word, Wordreference, Yahoo y Youtube.*
- Es el centro con mayor número de siglas: *ADSL, CD burner, CD recorder, CD-Rom, CPU, DVD, Gb, HTML, IT assistant, ISP, LCD screen, LED, MP3, MS-DOS, Mb, PC, PDF, RAM, RAM memory, ROM, Tb, TFT screen, TIC y URL.*

b) Edición en lengua francesaCentro 1. *Parties du corps humain*

- Plurales irregulares: se ha incluido *yeux* que es el plural irregular de *œil*.
- Vocablos marcados como vulgares en el diccionario: *bite, chatte y cul.*
- Gran parte de las palabras han sido escritas en plural por los informantes aunque han sido lematizadas en singular: *dents, doigts, mains, jambes*, etc.

Centro 2. *Les vêtements*

- Aparece una gran cantidad de palabras de origen extranjero, sobre todo vocablos procedentes del inglés pero muy usuales en francés: *anorak, jeans, polo, pull-over, pyjamas, string, slip, smoking sweater y t-shirt.*

Centro 3. *La nourriture et les boissons*

- Palabras de origen extranjero: *capuccino, carpaccio, choucroute, chewing-gum, fricadelle, hamburger, hot dog, macaron, paëlla, pizza, spaghetti, rhum, steak, whisky y yaourt.*
- Vocablos formados por dos o más palabras: *andive au jambon, café au lait, canard à l'orange, côte de porc, crème fraîche, eau gazeuse, eau minérale, jus de fruit, jus*

²⁰¹ Centro muy rico y variado debido a que en español también se utilizan palabras de origen inglés.

d'orange, menthe à l'eau, pain au chocolat, paté en croûte, petit pois, plateau de fromage, pomme de terre, poulet rôti, salade verte, steak haché y steak tartare.

Centro 4. *L'école*

- Marcas: *Tipp-Ex* y *Post-it* están presentes en los tres idiomas. Además, se ha añadido una marca de cinta adhesiva que ya se encuentra nominalizada y aparece en los diccionarios en francés: *scotch*.

Centro 5. *La ville*

- Topónimos: *Málaga* y *Lyon*.

Centro 6. *Moyens de transport*

- Abreviaturas: *(auto)bus*, *AVE*, *métro*, *moto(cyclette)*, *RER*, *tram(way)*, *TER*, *TGV* y *vélo*.
- Unificación de grafías: *skate-board* y *skateboard* se han escrito como *skateboard*; *mini-bus* y *minibus* como *minibus*.
- Palabras coloquiales para referirse a *voiture*: *bagnole* y *caisse*.

Centro 7. *Jeux et loisirs*

- Nombres propios: *Blokus*, *Buzz!*, *Facebook*, *Monopoly*, *Loto*, *Loto sportif*, *Pilates*, *Play Station*, *Pro evolution*, *Scrabble*, *Super Mario*, *Trivial* y *Wii*.
- Palabras provenientes del inglés (dos en su forma abreviada): *aquagym*, *baseball*, *basket(ball)*, *beach volley*, *cinéma*, *cricket*, *disco*, *foot(ball)*, *fitness*, *golf*, *handball*, *internet*, *paddle*, *poker*, *pub*, *rugby*, *shopping*, *ski*, *spinning*, *sport*, *squash*, *tennis*, *volleyball* y *weekend*.
- Juegos tradicionales: *la tomate* es un juego que consiste en regatear la pelota pasando el balón entre las piernas de su adversario; *cap ou pas cap?* es una frase hecha que quiere decir *capable ou pas capable?* se suele utilizar para empezar a jugar o para retar a alguien a hacer alto; *cache-cache* cuya traducción es escondite; *chat perché* es un tipo de juego tradicional similar al *se la queda*.

Centro 8. *Professions et métiers*

- Palabra sin traducción directa al español: *videur* que se refiere a la persona encargada de echar gente de un cabaret.

- Profesiones que los informantes han escrito en femenino: *actrice, boulangère, cassière, chanteuse, coiffeuse, femme de ménage, hôtesse, hôtesse de vol, infirmière, policière, pompière, professeure* y *servante*²⁰².

Centro 9. *Ordinateurs et Internet*

- Prevalencia de palabras anglófonas: *blog, chat, CD, clé USB, computer, DVD, email, hardware, hifi, keyboard, LCD, mouse, netbook, page web, pen drive, scanner, spam, site web, software, USB, web, webcam* y *wifi*.
- *Mél*: es una palabra coloquial para decir *courrier électronique, mail*.
- Nombres propios: *Adobe Photoshop, Busuu, CityVille, eMule, Facebook, Google, Hot Potatoes, Interplex, JClick, Les Sims, Linux, Microsoft Excel, Microsoft Office, Microsoft PowerPoint, (Microsoft) Word, Microsoft Works, MP4, MSN, PSP, Skype, Spotify, Tuenti, Twitter, Web 2.0, Wikipedia, Windows, Word, Wordreference, Yahoo* y *Youtube*.
- Abreviaturas: *CD, DVD, LCD, MP4, PSP* y *USB*.

3.5.3. Prueba del léxico especializado

En la edición de los datos de esta prueba se tuvieron en cuenta los dos tipos de análisis que se pretendían realizar. Desde un punto de vista cuantitativo, se contabilizaron las respuestas correctas, incorrectas y no contestadas. Se aplicaron diferentes fórmulas matemáticas para obtener los resultados de la prueba de identificación del léxico de tecnicismos que dieron lugar al Pident y al índice corregido de identificación de tecnicismos (ICIT). A las respuestas correctas se le asignó el valor 2, a las incorrectas el valor 1 y a las no contestadas el 0 y se procedió a contabilizar el número de respuestas dadas en cada caso. El Pident es la suma de las respuestas correctas y el ICIT el resultado de la suma de las respuestas correctas menos la resta de las respuestas incorrectas con un valor estadístico ponderado.

Desde un enfoque cualitativo se registraron las respuestas dadas para cada palabra de forma que no solo se obtuviese información sobre si el informante ha contestado bien o mal sino que también se pudiesen relacionar las respuestas de cada

²⁰² El masculino es irregular por lo que se ha añadido el masculino y el femenino como dos vocablos diferentes: *serviteur-servante*.

hablante con el centro de interés. Así, es posible detectar los centros con mayor y menor número de respuestas incorrectas e intentar buscar una explicación a este hecho.

La codificación se realizó incluyendo el número del centro de interés aportado por cada palabra. Por ejemplo, si el informante ha relacionado el vocablo proctólogo con el centro de interés número 8, *Profesiones y oficios*, se considera correcto y se le asigna el valor 2. No obstante, en el supuesto que el informante le haya otorgado el número 9, se codifica con el valor 1 y, además, se registra que se ha confundido con el centro *Ordenadores e internet*.

3.6. Conclusiones parciales

A lo largo de este capítulo, se han abordado los principios epistemológicos en los que se fundamenta el presente estudio. Tras realizar una reflexión sobre el paradigma cualitativo y positivista, se ha optado por incluir una metodología que integra ambos métodos según los objetivos de investigación. Por un lado, en los capítulos sucesivos se realizarán análisis cualitativos sobre la organización del lexicón mental de los informantes. Por otro lado, se abordarán cuáles son las variables que influyen en el léxico disponible y en la capacidad de identificación de tecnicismos desde el ámbito de la léxico-estadística.

La muestra de estudio está conformada por estudiantes de postgrado que cursan el Máster en Profesorado. De esta forma, es posible realizar una evaluación léxica del futuro profesorado, factor clave en la implantación de programas de bilingüismo. El acceso a los informantes se ha cuidado en detalle para lo que se ha establecido un documento de negociación que se ha entregado previamente a los agentes implicados. Se les ha informado de los objetivos de la investigación y de sus características. Además, se ha advertido la confidencialidad de los datos y se les ha facilitado los datos de contacto de la doctoranda al profesorado y al alumnado que deseaba tener alguna referencia complementaria.

En esta fase de la investigación, ha sido imprescindible el seguimiento a través de un diario en el que se anotaban todos los aspectos organizativos: especialidades en las que se podía realizar los cuestionarios, horario de tutorías del profesorado,

tamaños de los grupos, etc. Cabe destacar que el máster está dividido en catorce especialidades diferentes y se han conseguido los datos de más de la mitad de los informantes matriculados en diez especialidades. En cuanto a las cuatro restantes, no fue posible pasar las encuestas por diferentes motivos, sobre todo porque el alumnado tenía ya todo el contenido fijado de sus clases y el profesorado responsable de la asignatura no respondía a la petición de colaboración.

Los datos extralingüísticos se han recogido a través de un cuestionario con preguntas sociológicas, reticulares y relacionadas con el proceso de adquisición y aprendizaje de la lengua extranjera del informante. La información lingüística se ha extraído mediante la prueba de disponibilidad léxica y un test de identificación de tecnicismos con los mismos centros de interés en español y en lengua extranjera (inglés o francés). Estas áreas temáticas corresponden a ocho centros²⁰³ del estudio pionero de Gougenheim *et al.* (1964) y un centro de nueva incorporación, *Ordenadores e internet*. Cabe destacar que aunque los dieciséis centros tradicionales estén presentes en la mayoría de estudios de disponibilidad léxica, ya desde el año 1971, Mackey se planteaba su idoneidad:

Bien qu'elle reproduise les résultats de la première enquête française, cette étude témoin laisse à résoudre certaines questions de base, à savoir: Combien de centres d'intérêt sont communs à toute une population? Combien de réponses suffisent pour chaque centre? Quel âge et quel échantillon de personnes sont représentatifs de la langue commune? (Mackey 1971a: 13, *apud* Carcedo González, 1998b: 20).

Una vez recogidos los datos, se realizó un plan de codificación para el cuestionario sociológico que consiste en atribuir un valor numérico a cada respuesta e incorporarlo a una base de datos. Para la lematización del léxico disponible, se siguieron las recomendaciones de Samper Padilla (1998) seguidas en la mayoría de los trabajos herederos del PPHDL. Asimismo, se introdujeron criterios propios tales como la incorporación a los listados de los nombres propios y las asociaciones individuales. Esta inclusión permite realizar análisis más fidedignos sobre la relación entre palabras. Las peculiaridades de la edición han sido recogidas y clasificadas según el idioma y el centro de interés para orientar a futuros investigadores y servir de

²⁰³ 1) *Partes del cuerpo humano*; 2) *La ropa*; 3) *Comidas y bebidas*; 4) *La escuela: muebles y materiales*; 5) *La ciudad*; 6) *Medios de transporte*; 7) *Juegos y distracciones*; 8) *Profesiones y oficios*.

referencia en la lematización de algunos vocablos. En cuanto a la prueba de identificación del léxico especializado, las respuestas se clasificaron según fueran correctas, incorrectas o no contestadas²⁰⁴.

En síntesis, este estudio se sirve de la metodología de base de las investigaciones de disponibilidad que giran en torno al PPHDL pero también incorpora avances metodológicos como la prueba de identificación del léxico especializado creada empleada por Ávila Muñoz y Villena Ponsoda (2010) o la utilización del programa DispoGrafo (Echeverría *et al.*, 2008). Asimismo, incorpora una muestra de estudio original sobre la que se evalúa el léxico en lengua materna y lengua extranjera. Como apunta Moreno Fernández (1990: 38): «sólo así la comparación es posible, sólo así los pasos de una persona pueden ser de plena utilidad para otra, porque los avances se consiguen a mitad de camino entre la tradición y la originalidad».

²⁰⁴ El procesamiento matemático de este cuestionario así como de la prueba de léxico disponible se fundamentan en el apartado 2.5.4 y se ponen en práctica en forma de análisis en los capítulos 4, 5 y 6.

CAPÍTULO 4. ANÁLISIS GENERAL

- 4.1. Introducción**
- 4.2. Análisis cuantitativo de palabras y vocablos**
 - 4.2.1. Palabras y vocablos en español
 - 4.2.2. Palabras y vocablos en lengua extranjera
 - 4.2.3. Palabras y vocablos en lengua inglesa
 - 4.2.4. Palabras y vocablos en lengua francesa
- 4.3. Índice de densidad léxica e índice de cohesión**
 - 4.3.1. Índice de densidad léxica e índice de cohesión en español
 - 4.3.2. Índice de densidad léxica e índice de cohesión en inglés
 - 4.3.3. Índice de densidad léxica e índice de cohesión en francés
- 4.4. Promedio de palabras**
 - 4.4.1. Comparación intramuestral
 - 4.4.1.1. Amplitud léxica
 - 4.4.1.2. Índice corregido de amplitud léxica
 - 4.4.2. Comparación intermuestral
 - 4.4.2.1. Estudios en la provincia de Málaga
 - 4.4.2.2. Estudios en lengua extranjera
- 4.5. Análisis de los vocablos con mayor índice de disponibilidad**
 - 4.5.1. Comparación intramuestral
 - 4.5.2. Comparación intermuestral
 - 4.5.2.1. Investigaciones en español
 - 4.5.2.2. Investigaciones en inglés y francés
- 4.6. La situación ortográfica**
 - 4.6.1. Resultados del estudio actual
 - 4.6.2. Resultados comparativos intermuestrales
- 4.7. Incidencia de la variable ‘sexo’ en el léxico disponible**
 - 4.7.1. Centro 2: *La ropa*
 - 4.7.2. Centro 3: *Comidas y bebidas*
 - 4.7.3. Centro 4: *La escuela: muebles y materiales*
- 4.8. Conclusiones parciales**

4.1. Introducción

Este capítulo centra su atención en los datos generales relativos a la prueba de disponibilidad léxica: número total de palabras por prueba y por centro de interés, número de vocablos, índice de cohesión y densidad léxica. Esos datos se complementan con el análisis ortográfico y de una de las variables más estudiadas, el sexo. El objetivo es evaluar si el léxico disponible de estudiantes de postgrado ha evolucionado cuantitativa y cualitativamente a través de la corrección ortográfica. Asimismo, se pretende conocer si el léxico en lengua extranjera dista mucho del léxico en lengua materna.

En primer lugar, se analiza el número de palabras y vocablos de este estudio. Esta indagación se completa a través de la comparación del promedio de las lexías recogidas en esta investigación con respecto a encuestas previas. Se parte de la hipótesis que los listados en lengua materna son más productivos que los listados en lengua extranjera. En segundo lugar, se presentan los resultados del índice de densidad léxica y el índice de cohesión. En esta ocasión, no se realizan comparaciones con otros estudios por tratarse de dos índices directamente relacionados con el número de informantes. Se pretende conocer la estructura de los centros de interés para saber si guardan cierta relación con independencia de la lengua de la prueba.

En tercer lugar, se presentan los vocablos más disponibles con objeto de analizar la diferencia entre varias sintopías y entre el léxico aportado en la lengua materna y en la lengua extranjera de los informantes. La comparación intramuestral coteja los veinte vocablos más disponibles en español, inglés y francés y los presenta visualmente marcados para facilitar la lectura, además de incluir la distribución porcentual de vocablos coincidentes en los tres listados. La comparación intermuestral se divide según se trate de la lengua española, inglesa o francesa. En lengua española, los listados se parangonan con estudios de español como lengua extranjera (Carcedo González, 2000c; Samper Hernández, 2002) y otros estudios realizados en la provincia de Málaga a hablantes nativos (Ávila Muñoz, 2006; Ávila

Muñoz y Villena Ponsoda, 2010). En inglés, se han incluido la primera investigación en este idioma (Dimitrijević, 1969) y la realizada por Ferreira (2006) a estudiantes universitarios de Pedagogía en Inglés y Traducción en Lengua Extranjera y hablantes nativos. En francés, se ha comparado con el estudio pionero en el campo de la disponibilidad (Gougenheim *et al.*, 1964).

En cuarto lugar, se recogen las faltas de ortografía presentes en las encuestas según idiomas (español, inglés, francés) para realizar un análisis por centro de interés y tipo de error ortotipográfico para comprobar si la mayoría de faltas de ortografía se deben a errores de acentuación. A continuación, se comparan los resultados con otros estudios basados en encuestas de estudiantes preuniversitarios y universitarios para extraer conclusiones que desvelen el grado de incidencia de los estudios del informante en la corrección ortográfica de las encuestas.

Por último, se estudia el efecto del sociolecto femenino y masculino en el número de palabras y vocablos aportados por los informantes. Una vez detectados los centros de interés con mayor número de palabras exclusivas de un sexo u otro con diferencias significativas en el promedio total de palabras, se analizan con más detenimiento: *La ropa*, *Comidas y bebidas* y *La escuela: muebles y materiales*.

4.2. Análisis cuantitativo de palabras y vocablos

4.2.1. Palabras y vocablos en español

La distribución de palabras y vocablos en los diferentes centros varía de forma considerable. La Tabla 4.1 muestra el reparto de las unidades léxicas por centro de interés, la media y el rango de productividad. Los datos oscilan entre las 4493 unidades léxicas totales del centro más productivo, *Comidas y bebidas*, a 2641 palabras del centro clasificado con menor rango, *Medios de transporte*. La diferencia en la media por informantes entre un centro y otro asciende a más de 10 palabras, ya que el primer centro según rango cuenta con una media de 26,7 producciones por informante y el último tan solo 15,44.

Centros de interés	Palabras	Promedio	Rango	Vocablos	Promedio	Rango
01. Partes del cuerpo	4493	26,28	2	244	1,43	7
02. La ropa	3759	21,98	3	238	1,39	8
03. Comidas y bebidas	4565	26,7	1	530	3,10	2

04. La escuela: muebles y materiales	3718	21,74	4	392	2,29	6
05. La ciudad	3711	21,7	5	483	2,82	3
06. Medios de transporte	2641	15,44	9	190	1,11	9
07. Juegos y distracciones	3040	17,78	8	469	2,74	5
08. Profesiones y oficios	3600	21,05	6	478	2,80	4
09. Ordenadores e internet	3290	19,24	7	555	3,25	1
Total	32817	191,91		3579	20,93	

Tabla 4.1. Número total de palabras y vocablos y sus respectivos promedios y rangos de productividad en español según el centro de interés

Los informantes han escrito un total 32 817 palabras en los nueve centros de interés, lo que supone una media de 3646,33 lexías en total y 191,91 por centro de interés. Los cinco primeros centros de interés superan el promedio. Sin embargo, los cuatro siguientes ocupan las posiciones con menor rango y menor productividad.

En lo que concierne a la distribución porcentual de las palabras según centros de interés, existe una diferencia de más de un 5 % entre el centro más y menos productivo. El tercer centro cuenta con un 13,91 % de las palabras totales y el sexto con un 8,05 %. Los cinco primeros centros superan la media general de productividad del 11,11 % y suponen el 61,69 % de las palabras producidas por los informantes. Asimismo, se observa que los cuatro primeros centros completan más del 50 %.

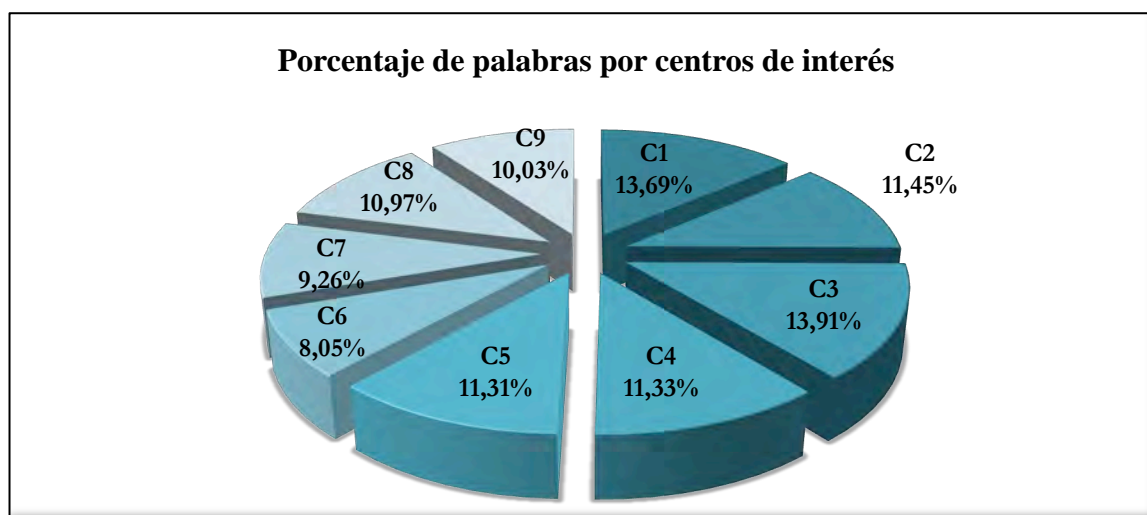


Gráfico 4.1. Distribución de las palabras por centros de interés

En cuanto al número de vocablos, los 171 informantes han aportado un total de 3579 unidades léxicas diferentes. Tanto el reparto por centros de interés, como el consiguiente rango difieren de la distribución del número de palabras. El centro con mayor número de vocablos es el referente a *Ordenadores e internet* con 555 palabras y,

con una diferencia de 365 vocablos se sitúa el centro con menor número, *Medios de transporte*, con 190 vocablos. Además, este último centro repite el puesto de último rango en el número de palabras.

El promedio de vocablos por centros de interés es de 397,67. Por encima de la media se encuentran los centros *Ordenadores e internet*, *La ciudad*, *Profesiones y oficios*, *Juegos y distracciones* y *Comidas y bebidas*. Por debajo de la media y por orden de productividad se situarían: *La escuela: muebles y materiales*, *Partes del cuerpo*, *La ropa* y *Medios de transporte*.

En el caso de la distribución porcentual, se constata que los cinco centros que superan la media suponen el 70,28 % del total de vocablos producidos por los informantes. Asimismo, se observa que el centro con más vocablos, *Ordenadores e internet*, con un 15,51 % casi triplica al centro con menos, *Medios de transporte*, con 5,31 %. Por tanto, la diferencia entre el número de vocablos por centros es mucho mayor que la presentada en cuanto al número de palabras. En el primer caso, el intervalo entre el centro más y menos productivo apenas era de un 5,86 % frente al 10,20 % en el segundo.

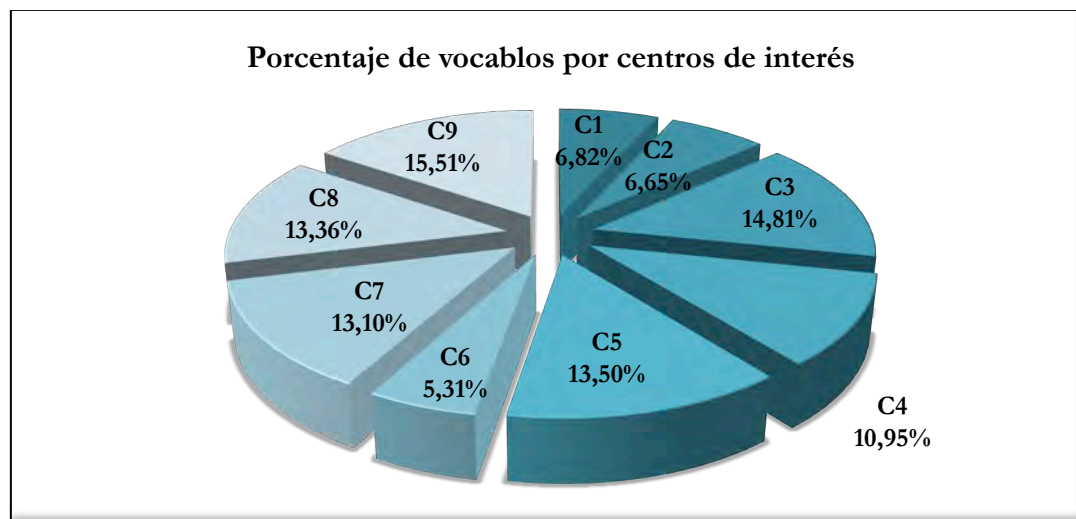


Gráfico 4.2. Distribución porcentual del número de vocablos por centros de interés

Como se aprecia en la comparación por rangos de la Tabla 4.1, solo coincide la posición número nueve del centro *Medios de transporte*. El rango superior en número de palabras lo ocupa el centro *Comidas y bebidas*, que también es el segundo en cuanto a número de vocablos. Sin embargo, el centro con mayor número de vocablos, *Ordenadores e internet*, se sitúa en séptimo lugar según la cantidad de palabras. Estos

primeros resultados arrojan datos exploratorios para conocer la configuración interna de los centros de interés.

4.2.2. Palabras y vocablos en lengua extranjera

El número total de palabras aportadas en lengua extranjera es considerablemente menor que en español, concretamente de 22 870 unidades léxicas y la media por centro de interés es de 2541,11 lexías. Se constata que solo cuatro centros superan la media. Por orden de rango, se sitúan los centros *Food and drink/ La nourriture et les boissons*, *Human body parts/ Parties du corps humain*, *The city/ La ville* y *Computer and internet/ Ordinateurs et internet*.

Centros de interés	Palabras	Promedio	Rango
01. Human body parts/ Parties du corps humain	2998	17,53	2
02. Clothes/ Les vêtements	2441	14,27	6
03. Food and drink/ La nourriture et les boissons	3346	19,57	1
04. The school (furniture and school material)/ L'école (meubles et matériel)	2443	14,29	5
05. The city/ Laville	2733	15,98	3
06. Means of transport/ Moyens de transport	1924	11,25	9
07. Games and entertainments/ Jeux et loisirs	2191	12,81	8
08. Professions and jobs/ Professions et métiers	2235	13,02	7
09. Computer and Internet/ Ordinateurs et Internet	2567	15,01	4
Total	22870	2541,11	

Tabla 4.2. Número total de palabras, promedio y rango de productividad en lengua extranjera

Desde el punto de vista de la formación pedagógica del futuro profesorado, el centro referido a la escuela y a las nuevas tecnologías debería contar con un mayor número de palabras por ser el vocabulario más empleado en clase. Sin embargo, los centros más productivos son los referidos a realidades útiles a la hora de realizar un viaje al extranjero como la comida, el cuerpo y la ciudad. Aunque *a priori* el cuerpo humano parece no guardar relación con los viajes, en todos los manuales y guías de viaje suelen incluir el tema por cuestiones relacionadas con la asistencia médica en caso de necesidad.

Dado que el alumnado que ha realizado las pruebas de disponibilidad ha debido acreditar previamente un nivel B1 según el Marco Común Europeo de Referencia para las Lenguas (MCERL), se hace patente la necesidad de comprobar si el conocimiento adquirido es similar a los contenidos mínimos fijados para ese nivel. En el cuadro referido a los niveles del MCERL se especifica que el estudiante que adquiere un nivel B1:

Es capaz de comprender los puntos principales de textos claros y en lengua estándar si tratan sobre cuestiones que le son conocidas, ya sea en situaciones de trabajo, de estudio o de ocio.

Sabe desenvolverse en la mayor parte de las situaciones que pueden surgir durante un viaje por zonas donde se utiliza la lengua.

Es capaz de producir textos sencillos y coherentes sobre temas que le son familiares o en los que tiene un interés personal.

Puede describir experiencias, acontecimientos, deseos y aspiraciones, así como justificar brevemente sus opiniones o explicar sus planes (Consejo de Europa, 2002: 26).

Por lo tanto, el hecho de que los centros de interés en lengua extranjera más productivos ayuden al estudiante a expresarse en viajes, casa perfectamente con los contenidos del nivel B1 del MCERL. Asimismo, el apartado titulado «La coherencia del contenido de los niveles comunes de referencia» incluye una breve descripción para cada uno, relacionando el nivel B1 directamente con la competencia lingüística de un viajero.

El nivel B1 refleja la especificación del nivel umbral para un viajero que visita un país extranjero, y tiene quizás dos características principales. La primera es la capacidad de mantener una interacción y de hacerse entender en una variedad de situaciones; [...] La segunda característica es la capacidad de saber cómo enfrentar de forma flexible problemas cotidianos, como, por ejemplo: se enfrenta a situaciones menos corrientes en el transporte público, a las típicas situaciones que suelen surgir cuando se realizan reservas para viajes a través de una agencia o cuando se está viajando (Consejo de Europa, 2002: 37).

Tal y como se puede observar en la siguiente figura, la diferencia entre los centros más y menos fructíferos es considerable, produciéndose una desigualdad de más de 1000 palabras entre el tercer centro, *Food and drink/La nourriture et les boissons*, y el sexto, *Means of transport/Moyens de transport*. Cabe resaltar que la distribución de palabras en lengua española es muy similar, por lo que se muestra la influencia de temas más abiertos, como *Comidas y bebidas*, y el carácter más restrictivo de *Medios de transporte*.

Respecto a la media porcentual, los cuatro centros de interés que superan el promedio de 11,11 %, alcanzan el 50,91 % del total de palabras producidas. La diferencia entre el centro de interés con más palabras, *Comidas y bebidas*, con un 14,63 %, y el centro con menos, *Medios de transporte*, con 8,41 %, es de 6,22 %. Esa escala es muy similar a la producida en lengua española entre los mismos centros, concretamente entre el 13,91 % y el 8,05 %. En ambos casos, coinciden el centro más y menos productivo pero difieren en el resto de la clasificación por rangos.

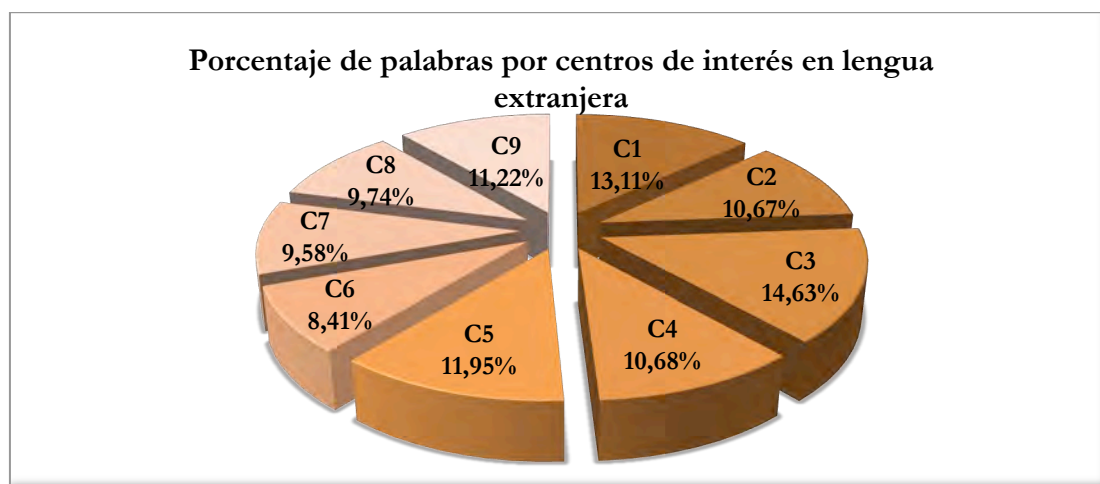


Gráfico 4.3. Distribución porcentual de las palabras por centros de interés en lengua extranjera

A continuación, se presentan los resultados de vocablos en lengua extranjera divididos según idiomas: inglés y francés. Se trata de dos idiomas diferentes con lo que se podrían traducir literalmente las palabras en los casos que fuese posible o considerar los vocablos en cada idioma por separado. En este caso, se daría la tesitura de contar doblemente palabras de un mismo origen, significante y significado en sendos idiomas, como podría ser *whisky* o *pizza*. Asimismo, el número de informantes influye en el número de vocablos por lo que se ha optado por analizar cada idioma por separado sin realizar comparaciones con otros estudios.

4.2.3. Palabras y vocablos en lengua inglesa

Los 150 informantes han escrito un total de 19 768 palabras y una media por centros de interés de 131,79 palabras. Al igual que ocurre en lengua española y en lengua extranjera, el centro más productivo es el tercero, *The school: furniture and school material*, con 2894 palabras, y el menos fructífero el sexto con 1675.

Centros de interés	Palabras	Promedio	Rango	Vocablos	Promedio	Rango
01. Human body parts	2571	17,14	2	115	0,67	9
02. Clothes	2076	13,84	6	151	0,88	8
03. Food and drink	2894	19,29	1	287	1,68	4
04. The school: furniture and school material	2120	14,13	5	244	1,43	6
05. The city	2344	15,63	3	343	2,01	2
06. Means of transport	1675	11,17	9	165	0,96	7
07. Games and entertainments	1909	12,73	8	310	1,81	3
08. Professions and Jobs	1928	12,85	7	280	1,64	5
09. Computer and Internet	2251	15,01	4	460	2,69	1
Total	19768	131,79		2355	13,77	

Tabla 4.3. Número de palabras y vocablos y sus respectivos promedios y rangos de productividad según el centro de interés

Los centros que superan el promedio de palabras, es decir de 2196,44 léxias, son por orden de rango el tercero, *Food and drink*; el primero, *Human body parts*; el quinto *The city* y el noveno, *Computer and Internet*. A diferencia de los resultados en español, el centro *Computer and Internet* supera la media dejando constancia del papel predominante que juega la lengua inglesa en las nuevas tecnologías.

Los cuatro centros más productivos superan la media del total de palabras suponiendo el 50,89 % de la muestra. La diferencia porcentual entre los centros más y menos productivos oscila entre el 13,01 % y el 8,47 %, porcentajes muy similares a los presentados previamente en lengua extranjera.

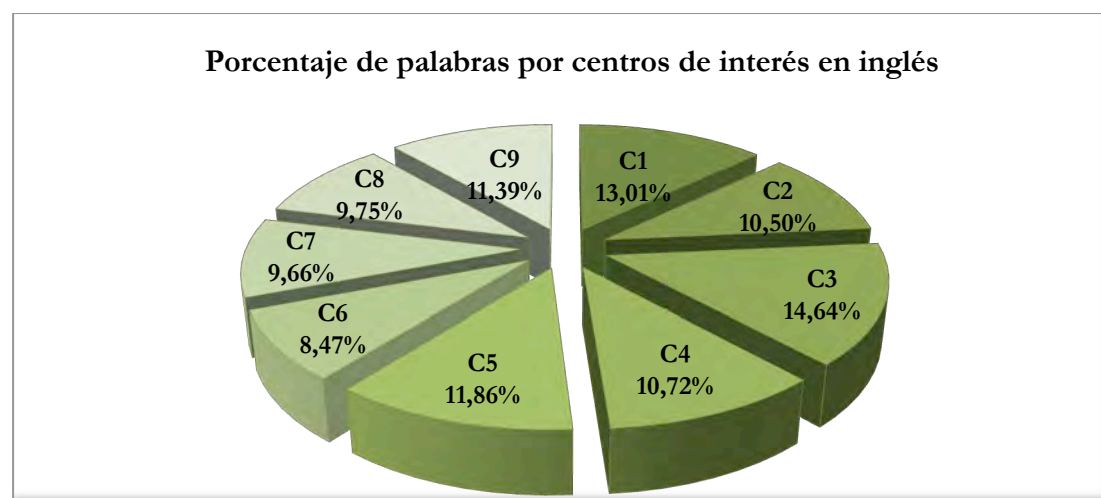


Gráfico 4.4. Distribución porcentual de las palabras por centros de interés en inglés

En el caso de los vocablos, también el centro más productivo ha sido *Computer and Internet* con 460 unidades léxicas diferentes. El centro con menos riqueza léxica, *Human body parts*, cuenta con 115 vocablos diferentes, con lo que se produce una diferencia sustancial de 345 vocablos entre ambos. La media de vocablos por centros

es de 261,67 aunque su distribución es muy desigual. Así, el centro más rico en léxico, *Computer and Internet*, supera ampliamente el promedio. Por orden de rango y superando la media, le siguen el quinto, séptimo, el tercer y el octavo centro de interés. De nuevo, las áreas temáticas relacionadas con los ordenadores, la ciudad, la comida y las distracciones son las más productivas, lo cual encaja con el léxico de un estudiante que ha finalizado un nivel B1.

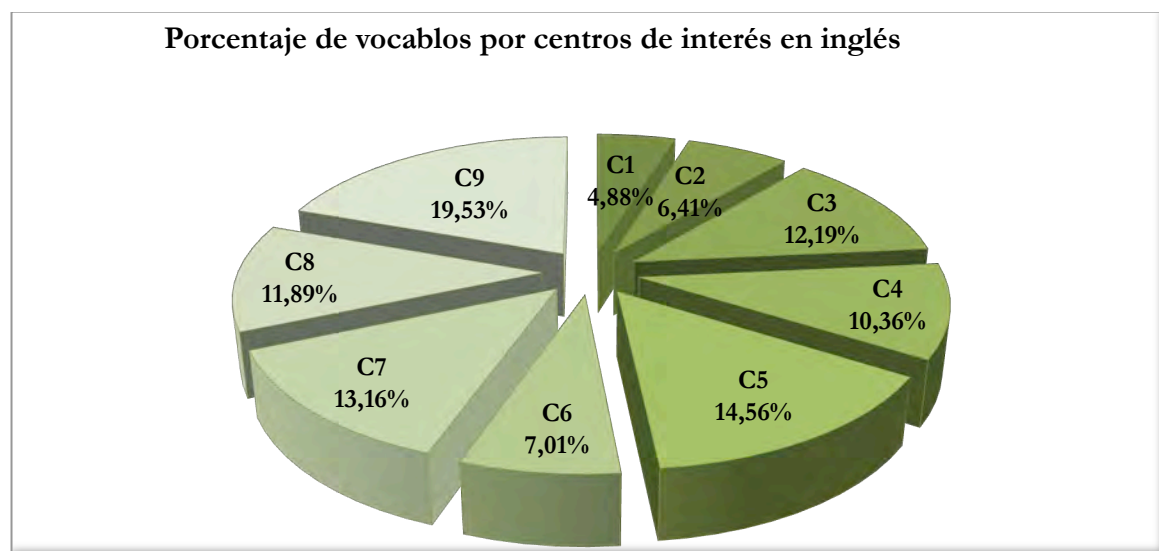


Gráfico 4.5. Distribución porcentual de los vocablos por centros de interés en inglés

Los cinco centros que superan la media suponen el 71,34 % de todos los vocablos en lengua inglesa. El Gráfico 4.1 muestra una disparidad entre el centro menos productivo, *Human body parts*, y el centro más productivo, *Computer and internet*, exactamente los datos se cuatriplican pasando de un 4,88 % a un 19,53 %.

4.2.4. Palabras y vocablos en lengua francesa

Los 21 informantes han aportado un total de 3102 palabras y una media de 147,71 palabras por centro de interés. *La nourriture et les boissons*, es el centro que ocupa el primer rango y, al igual que ocurre en lengua española, el centro referido a los transportes ocupa la última posición. Aunque el alumno posea un nivel en el que debe saber desenvolverse en los viajes, el hecho de que el centro menos productivo sea *Moyens de transport* no sería significativo puesto que en español muestra un comportamiento similar.

Centros de interés	Palabras	Promedio	Rango	Vocablos	Promedio	Rango
01. Parties du corps humain	427	20,33	2	78	0,46	8
02. Les vêtements	365	17,38	4	80	0,47	7
03. La nourriture et les boissons	452	21,52	1	166	0,97	1
04. L'école: meubles et matériel	323	15,38	5	113	0,66	6
05. La ville	389	18,52	3	135	0,79	2
06. Moyens de transport	249	11,86	9	53	0,31	9
07. Jeux et loisirs	282	13,43	8	125	0,73	4
08. Professions et métiers	299	14,24	7	119	0,70	5
09. Ordinateurs et Internet	316	15,05	6	127	0,74	3
Total	3102	147,71		996	5,82	

Tabla 4.4. Número de vocablos y rango por centro de interés en francés

Tal y como se puede apreciar, cuatro centros superan el promedio de 344,67 palabras. El centro que ocupa el primer rango es el tercero, *La nourriture et les boissons*, seguido de *Parties du corps humain*, *La ville* y *Les vêtements*. El resto de centros están por debajo del promedio aunque no muy alejados puesto que hay una distribución más equitativa de las palabras que en lengua española o inglés.

Los cuatro centros que superan la media porcentual, es decir los tres primeros y el quinto centro, suponen el 52,64 % del total de palabras producidas en francés. La distancia entre el centro uno y el seis se muestra un poco más acentuada que en inglés, siendo de 13,77 % frente a 8,03 %.

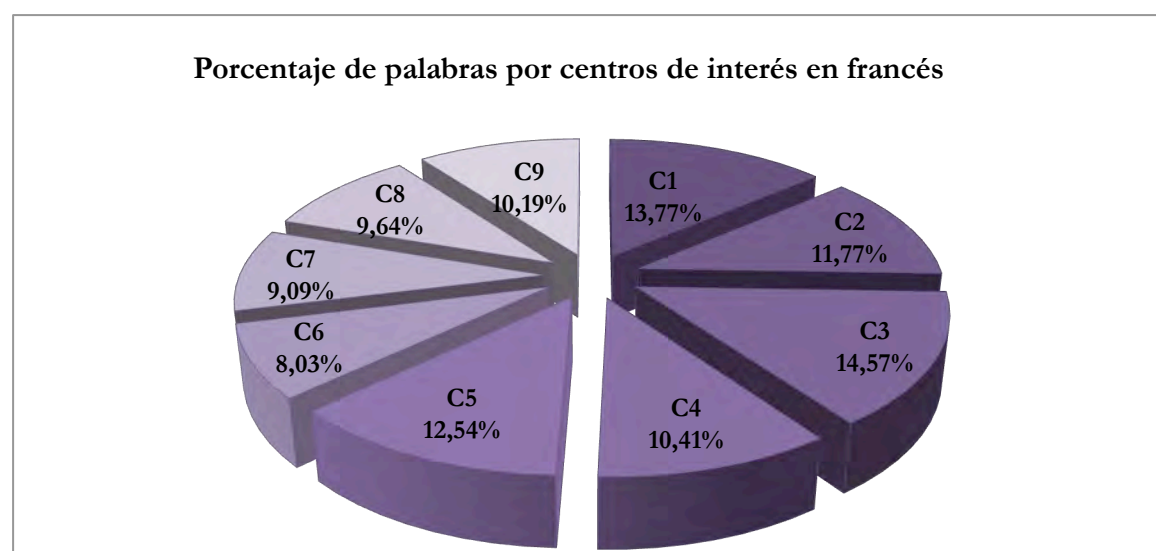


Gráfico 4.6. Distribución porcentual de las palabras por centros de interés en francés

En cuanto al número de vocablos, el total suma 996 en todos los centros dando lugar a una media es de 110,67 por centro de interés. En sintonía con los

resultados de palabras por centros, el tercero, *La nourriture et les boissons*, ocupa el primer rango con un total de 166 vocablos diferentes. En el lado opuesto se encuentra el sexto centro, *Moyens de transport*, con 53 vocablos.

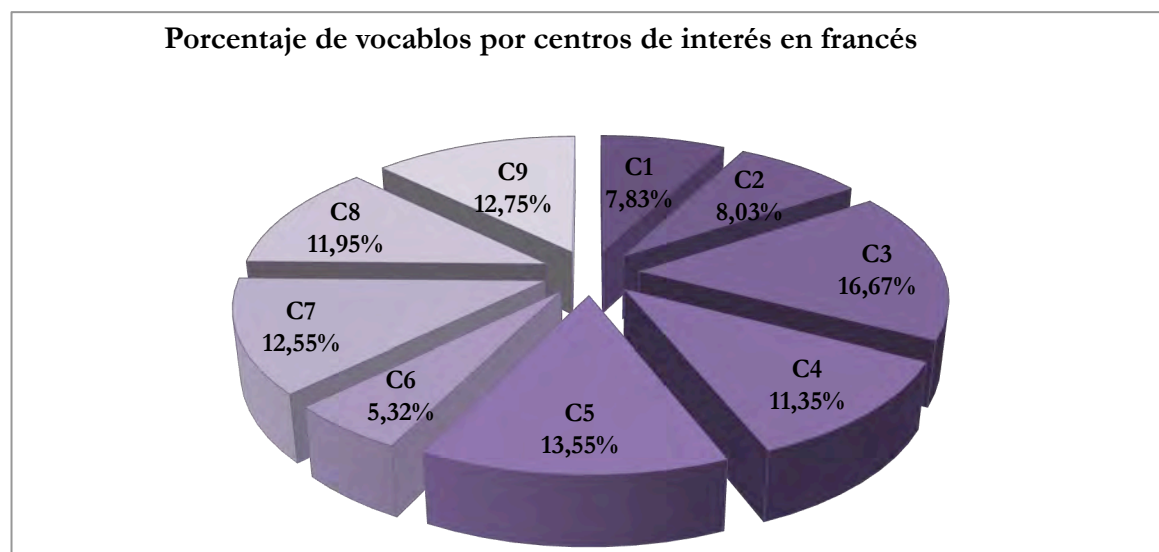


Gráfico 4.7. Distribución porcentual de los vocablos por centros de interés en francés

Al contrario que en el caso de la distribución de palabras, seis centros sobrepasan la media. Esos centros suponen 78,15 % de los vocablos en francés. Se observa que el tercer centro triplica el porcentaje de vocablos recogidos en el sexto centro. Se registra una gran diferencia entre el centro con más y menos vocablos de 103 palabras. Esa escala se produce en inglés y español pero de forma menos acentuada. En los tres casos idiomas el centro con menos número de vocablos y palabras es el sexto, *Moyens de transport*.

4.3. Índice de densidad léxica e índice de cohesión

En los apartados anteriores, se observa que el número de palabras y vocablos difiere de un centro a otro de forma considerable. Por un lado, se encuentran áreas temáticas heterogéneas con una gran cantidad de vocablos diferentes pero con un menor número de palabras. Por otro lado, existen centros compactos en los que el número de vocablos es más reducido y el número de palabras más elevado. Para conocer el grado de homogeneidad de cada centro, se aplica el índice de densidad léxica y el índice de cohesión.

El índice de densidad léxica fue empleado por primera vez por Alba Ovalle (1995a) en un estudio sobre disponibilidad léxica en la República Dominicana. El objetivo del autor era conocer el «mayor o menor grado de coincidencia en las respuestas» (Alba Ovalle, 1995a: 19). Para calcularlo, es necesario dividir en cada centro el total de palabras por el número de vocablos.

En cuanto al índice de cohesión, nace de la mano de los investigadores chilenos Mena Osorio en 1986 y Echeverría *et al.* en 1987 para quienes «mide el grado de coincidencia en las respuestas de los sujetos para un mismo centro» (Echeverría *et al.*, 1987: 68). Asimismo, Echeverría (1991) explica cómo aplicar este índice y cómo interpretarlo:

Si dividimos $\bar{X}R^{205}$ por NDP^{206} obtendremos un valor 1 siempre y cuando todos los sujetos hayan respondido las mismas palabras en su encuesta léxica, cosa por lo demás quimérica. Por el contrario, cuanto más variadas sean las respuestas, menor será el índice de cohesión; es decir, la dispersión es mayor. De este modo nuestro índice de cohesión informa si nuestro centro de interés es cerrado o compacto, o más bien abierto o difuso (Echeverría, 1991: 62).

Autores como Bartol Hernández (2004: 51) y Hernández Muñoz (2006: 311) dejan patente que el número de informantes condiciona el número de vocablos y, por ende, afecta directamente al cálculo de ambos índices. Asimismo, Bartol Hernández, (2004: 51) recuerda que «se debe tener en cuenta que el índice de cohesión está condicionado por el número de vocablos y por consiguiente, por las normas de edición». Siguiendo las recomendaciones anteriores, no se realizarán comparaciones entre las diferentes pruebas de este estudio ni con otros estudios por contar con un número diferente de informantes.

4.3.1. Índice de densidad léxica e índice de cohesión en español

La interpretación de ambos índices sigue un mismo patrón puesto que conforme aumenta la puntuación en ambos índices, mayor es la homogeneidad de la muestra. Los resultados del índice de cohesión revelan más uniformidad según se aproxime al valor 1, lo cual supondría que el 100 % de los informantes han contestado las mismas palabras. Se observa que el centro más cohesionado es el

²⁰⁵ Media de palabras por informante.

²⁰⁶ Número de palabras diferentes, es decir, vocablos.

primero, *Partes del cuerpo humano*, con un índice de 0,108. Por el contrario, el centro más disperso es el noveno, *Ordenadores e internet*, con 0,035. Ese mismo rango de puestos se repite en el índice de densidad léxica, donde el primer centro cuenta con 18,41 palabras y el noveno con 5,93.

CI	Palabras	Rango	Vocablos	Rango	Densidad léxica	Índice de cohesión	Rango
C1	4493	2	244	7	18,41	0,108	1
C2	3759	3	238	8	15,79	0,092	2
C3	4565	1	530	2	8,61	0,05	5
C4	3718	4	392	6	9,48	0,055	4
C5	3711	5	483	3	7,68	0,045	6
C6	2641	9	190	9	13,9	0,081	3
C7	3040	8	469	5	6,48	0,038	8
C8	3600	6	478	4	7,53	0,044	7
C9	3290	7	555	1	5,93	0,035	9
Media	2646,33		397,67		10,43	0,061	

Tabla 4.5. Densidad léxica e índice de cohesión en lengua española

Los centros de interés que superan la media de cohesión, es decir del 0,061, son el primero, *Partes del cuerpo humano*; el segundo, *La ropa*; y el sexto, *Medios de transporte*. Aunque el centro sexto ocupe el último rango en cuanto a productividad y número de vocablos, es un centro más cerrado donde la mayoría de las palabras aportadas por los informantes se repiten.

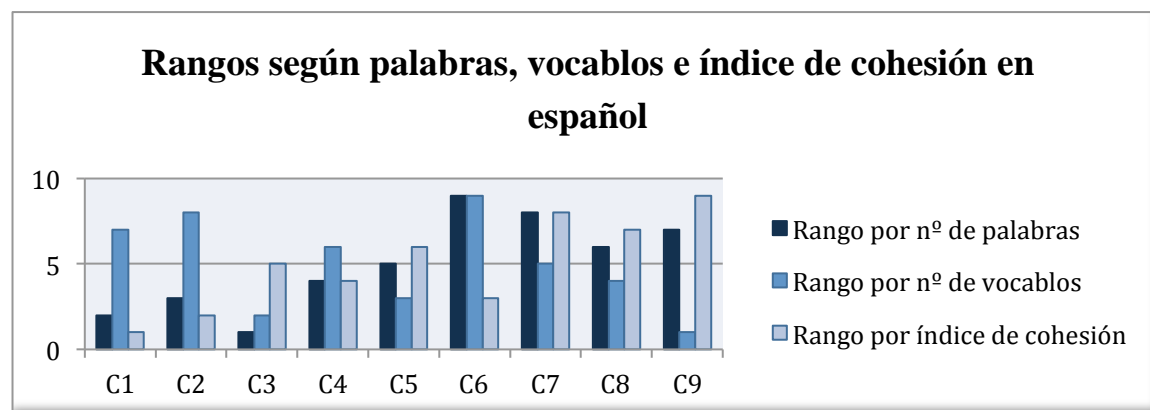


Gráfico 4.8. Rangos según palabras, vocablos e índice de cohesión en español

El Gráfico 4.8 presenta el rango que ocupa cada centro de interés según el número de palabras, vocablos y el índice de cohesión. Se observa que el primer y el segundo centro se sitúan en el segundo y tercer rango por número de palabras pero uno de los últimos rangos según el número de vocablos. Por tanto, ambos ocupan los

primeros rangos según el índice de cohesión indicando la homogeneidad que presentan sus léxicos.

4.3.2. Índice de densidad léxica e índice de cohesión en inglés

Los resultados en inglés muestran una gran escala entre el centro más y menos cohesionado. El centro más compacto es el primero, *Human body parts*, con una densidad de 22,36, seguido del centro dos, *Clothes*, con 13,75. Al igual que en los resultados en español, el centro menos cohesionado es el noveno, *Computer and Internet*, ya que sobre todo en inglés es un centro que cuenta con muchos vocablos y no tantas palabras en comparación con otros listados. En cuanto al índice de cohesión, la mayoría de centros presenta unos resultados en torno al promedio, 0,065, a excepción del primer centro.

CI	Palabras	Media palabras	Rango	Vocablos	Rango	Densidad léxica	Índice de cohesión	Rango
C1	2571	17,14	2	115	9	22,36	0,132	1
C2	2076	13,84	6	151	8	13,75	0,092	2
C3	2894	19,29	1	287	4	10,08	0,067	4
C4	2120	14,13	5	244	6	8,69	0,058	5
C5	2344	15,63	3	343	2	6,83	0,046	7
C6	1675	11,17	9	165	7	10,15	0,068	3
C7	1909	12,73	8	310	3	6,16	0,041	8
C8	1928	12,85	7	280	5	6,89	0,046	6
C9	2251	15,01	4	460	1	4,89	0,033	9

Gráfico 4.9. Palabras, vocablos, densidad léxica e índice de cohesión en inglés

Los centros de interés que ocupan los primeros puestos del rango en cuanto a número de palabras y de vocablos son los que presentan un índice de cohesión menor, por lo que ocupan un número más alto en el tercer rango. Por ejemplo, el centro nueve, *Computer and Internet*, presenta el primer rango en el número de vocablos y el último en cuanto a índice de cohesión ya que además es un centro poco productivo que se sitúa en el séptimo lugar en el rango por número de palabras.

En el caso opuesto, estaría el primer centro que ocupa el puesto uno en el rango del índice de cohesión pero el séptimo en cuanto al número de vocablos. Además, es el segundo centro más productivo, con lo que utilizando menos vocablos que en otros centros, los informantes escribieron muchas palabras, de ahí que sea el mayor en cuanto al índice de cohesión.

4.3.3. Índice de densidad léxica e índice de cohesión en francés

Al igual que ocurre en inglés y español, el centro más cohesionado es el primero, *Parties du corps humain*, por tratarse de uno de los centros con más palabras aportadas pero con menos vocablos diferentes. No obstante, los datos difieren en cuanto al centro menos uniforme ya que en este caso corresponde al séptimo, *Jeux et loisirs*, con un índice de 0,107.

CI	Palabras	Media palabras	Rango	Vocablos	Rango	Densidad léxica	Índice de cohesión	Rango
C1	427	20,33	2	78	8	5,47	0,261	1
C2	365	17,38	4	80	7	4,56	0,217	3
C3	452	21,52	1	166	1	2,72	0,130	6
C4	323	15,38	5	113	6	2,86	0,136	5
C5	389	18,52	3	135	2	2,88	0,137	4
C6	249	11,86	9	53	9	4,70	0,224	2
C7	282	13,43	8	125	4	2,26	0,107	9
C8	299	14,24	7	119	5	2,51	0,120	7
C9	316	15,05	6	127	3	2,49	0,118	8

Tabla 4.6. Palabras, Vocablos, densidad léxica e índice de cohesión en francés

Tan solo tres centros superan la media del índice de cohesión de 0,161, concretamente los mismos centros que en español. Cabe mencionar que ese promedio es bastante superior en francés que en inglés (0,065) o español (0,061) ya que el coeficiente se ve alterado por el menor número de informantes.

El tercer centro, *La nourriture et les boissons*, ocupa el primer puesto en el rango de número de palabras y vocablos. No obstante, el gran número de vocablos aportados hace que el centro no esté entre los primeros según cohesión. El siguiente en número de palabras es el primer centro, *Parties du corps humain*, pero como no ha aportado tantos vocablos diferentes es el centro con mayor índice de cohesión. Esto supondría que en los tres idiomas los informantes han escrito palabras cuya asociación es más directa en el centro uno y, por tanto, esas asociaciones han sido compartidas por la mayoría de ellos.

4.4. Promedio de palabras

La media de palabras por informante es un dato que no se ve alterado por el número de pruebas sino por el número de palabras que contienen cada una de ellas, lo que hace posible realizar una comparación con los resultados de otros estudios de

disponibilidad. En cambio, no se cotejan los datos relacionados con los vocablos o el índice de cohesión por verse directamente influenciado por el número de informantes.

En primer lugar, se tienen en cuenta las tres pruebas realizadas en el presente estudio correspondientes a los idiomas español, francés e inglés. Se presupone que los informantes han obtenido unos resultados superiores sobre el caudal léxico en español, por tratarse de la lengua materna; en francés, por provenir de la misma familia latina; y, por último en inglés. Esta comparación se realiza en base a dos parámetros: índice de amplitud léxica e índice corregido de amplitud léxica. El primero se refiere al número total de palabras, mientras que el segundo pondera las respuestas de los informantes según se trate de palabras nucleares ($N1 = 1$), asociaciones colectivas ($N5 = 0,5$), derivaciones ($N25 = 0,25$) y asociaciones individuales ($N0 = 0$).

A continuación, se toman como referencia dos investigaciones previas en la provincia de Málaga, permitiendo el parangón entre universos de población generacional y socioeconómica diferentes. Asimismo, se ha incluido el análisis comparativo con otras investigaciones centradas en el léxico en lengua extranjera (español, inglés y francés).

4.4.1. Comparación intramuestral

4.4.1.1. Amplitud léxica

El número total de palabras producidas en lengua española es 32817 y dando lugar a una media por informante 191,91 unidades léxicas. En lengua extranjera ambas cifras se reducen a 22870 palabras en total y 133,74 de media. En el Gráfico 4.10 esa diferencia notoria se observa entre el promedio de palabras en lengua española y en lengua extranjera por informante y centros, concretamente del 21,31 al 14,86. A pesar de ello, la distribución por centros de interés es muy similar, por lo que la curva ascendente y descendente en cada tema se repite excepto en el último centro de interés, referente a *Ordenadores e internet* cuyas medias se aproximan ligeramente.

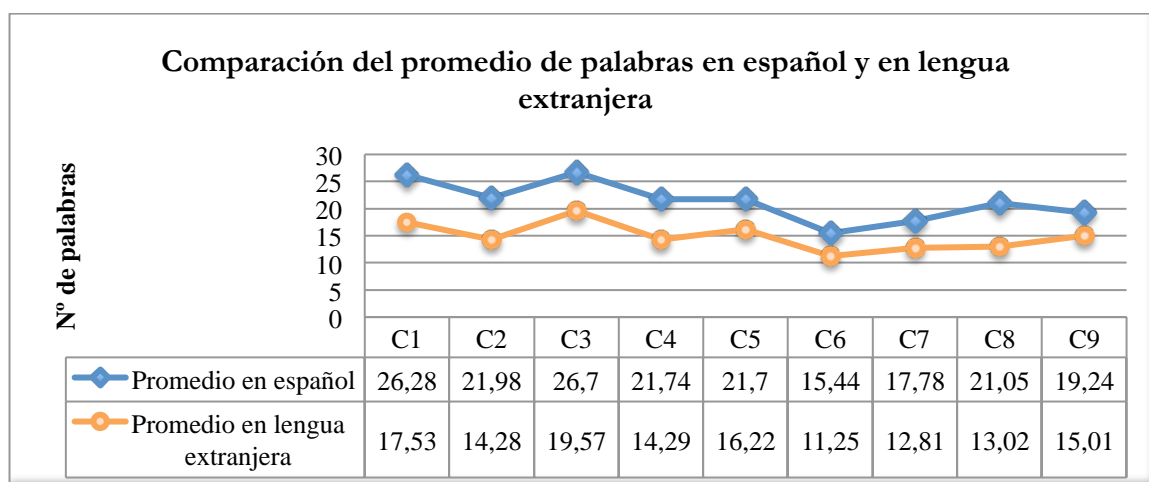


Gráfico 4.10. Comparación del promedio de palabras por centros de interés en español y en lengua extranjera

En cuanto a la comparación entre los dos idiomas extranjeros, se constata que los resultados en lengua francesa son más productivos que en lengua inglesa. El promedio de palabras por informante y centro de interés en inglés es de 14,64 mientras que en francés es de 16,41. En lengua extranjera la distancia entre los valores en inglés y francés se acortan, con lo que en algunos centros las líneas casi se unen, como es el caso del noveno, *Computer and Internet/Ordinateurs et Internet*, cuyos resultados son 15,01 y 15,05 palabras. En todos los centros el promedio de palabras es mayor en francés que en inglés y se repite la curva de distribución de datos presentes en el Gráfico 4.9.

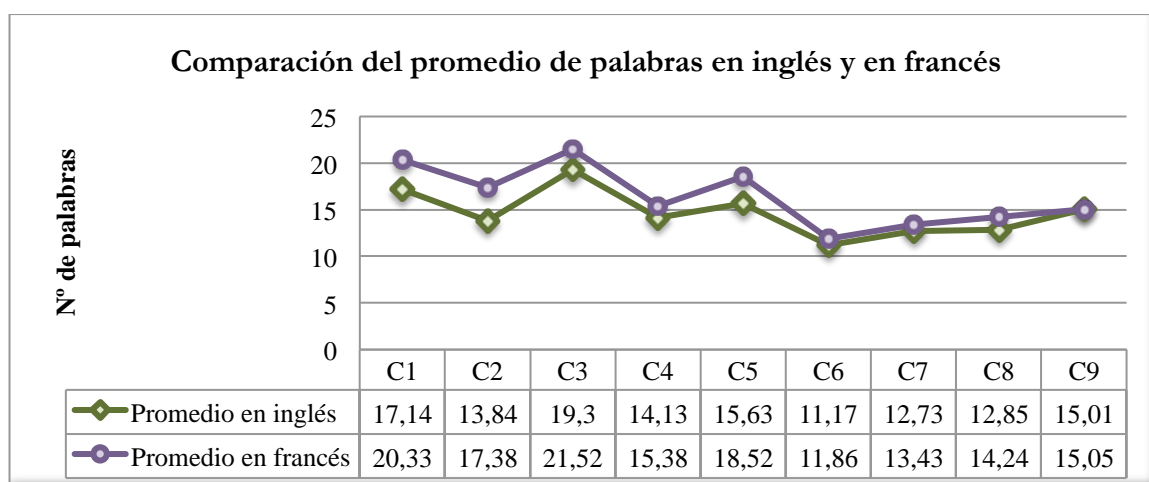


Gráfico 4.11. Comparación del promedio de palabras por centros de interés en inglés y francés

La Tabla 4.7 recoge el promedio de palabras y rango por centros de interés en español, lengua extranjera, inglés y francés. Se han destacado los centros de interés cuyos rangos coinciden en más de un idioma. De esta forma, se observa que el primer centro, *Human body parts/Parties du corps humain*, ocupa el rango dos; el tercero,

Food and drink/La nourriture et les boissons el puesto uno; el sexto centro, *Means of transport/Moyens de transport* el último puesto; y el séptimo centro, *Games and entertainments/Jeux et loisirs*, el puesto ocho.

CI	Media en español	Rango en español	Media en lengua extranjera	Rango en lengua extranjera	Media en inglés	Rango en inglés	Media en francés	Rango en francés
C1	26,28	2	17,53	2	17,14	2	20,33	2
C2	21,99	3	14,27	6	13,84	6	17,38	4
C3	26,70	1	19,57	1	19,29	1	21,52	1
C4	21,74	4	14,29	5	14,13	5	15,38	5
C5	21,70	5	15,98	3	15,63	3	18,52	3
C6	15,44	9	11,25	9	11,17	9	11,86	9
C7	17,78	8	12,81	8	12,73	8	13,43	8
C8	21,05	6	13,02	7	12,85	7	14,24	7
C9	19,24	7	15,01	4	15,01	4	15,05	6

Tabla 4.7. Promedio de palabras y rango por centros de interés en español, lengua extranjera, inglés y francés

En lengua extranjera, inglés y francés además coinciden el centro cuatro, *The school/L'école*, que ocupa el quinto puesto; y el centro cinco, *The city/La ville* en el tercer puesto. En inglés y en lengua extranjera²⁰⁷ repite posición el centro nueve, *Computer and Internet/Ordinateurs and Internet*, con el cuarto puesto en el rango. Las similitudes entre los resultados en lengua extranjera, por un lado, y en inglés y francés, por otro lado, se deben por la paridad en la distribución de palabras en inglés y francés, siendo siempre superior la media en francés.

4.4.1.2. Índice corregido de amplitud léxica

El índice de amplitud léxica se refiere al compendio de todas las palabras aportadas por un informante. En los listados del léxico disponible del estudio de Ávila Muñoz y Villena Ponsoda (2010) se recogen todas las unidades léxicas de forma que cuantas más palabras aporte un informante mayor será el resultado de este índice con independencia de su relación con el estímulo o tema dado. Para ponderar los resultados los investigadores asignaron valores diferentes a las palabras según su pertenencia al centro de interés distinguiendo cuatro tipo de unidades léxicas: palabras nucleares (N1) a las que se les asignó un valor de 1; asociaciones denotacionales (N5) con un valor de 0,5; derivaciones (N25) con un valor de 0,25 y

²⁰⁷ Los datos en lengua extranjera se refieren a la suma de palabras en inglés y francés.

asociaciones individuales (N0) con un valor de 0 (Ávila Muñoz y Villena Ponsoda, 2010). En primer lugar, se agrupan las palabras por tipos y a continuación se contabilizan todas las palabras de un mismo grupo y se dividen por la puntuación dada. La suma resultante de los cuatro conjuntos da lugar al índice corregido de amplitud léxica²⁰⁸.

En las Tabla 4.8 y Tabla 4.9 se pueden observar los datos de los diez informantes con mayor ICoAL en español y en lengua extranjera. Se observa que de diez sujetos en cada tabla, cuatro de ellos coinciden por ser los que mayor ICoAL muestran en español y en lengua extranjera. De esta forma se corrobora la hipótesis que apunta que el caudal léxico de un idioma y otro están relacionados de forma que cuanto mayor es el número de palabras en lengua extranjera, mayor será en lengua española. No obstante, se predice que la relación inversa está supeditada por la influencia de otros factores como las variables relacionadas con el aprendizaje de una lengua extranjera, razón por la cual los informantes con mayor léxico en lengua española aportan mayor número de palabras en francés o inglés si han realizado algún tipo de estudio previo en lengua extranjera.

Diez primeros informantes según el ICoAL en lengua extranjera												
Sujeto	P0 (LE)	P25 (LE)	P5 (LE)	P1 (LE)	IAL (LE)	ICoAL (LE)	P0	P25	P5	P1	IAL	ICoAL
24	5	4	14	213	236	221	2	10	2	196	210	199,5
38	3	2	7	211	223	215	0	5	12	197	214	204,25
16	10	14	23	188	235	203	2	19	28	251	300	269,75
8	6	7	13	192	218	200,25	4	6	21	235	266	247
11	3	8	11	189	211	196,5	7	6	9	157	179	163
21	2	15	9	186	212	194,25	1	6	19	268	294	279
20	5	8	5	186	204	190,5	0	13	5	211	229	216,75
15	2	3	4	184	193	186,75	6	7	8	252	273	257,75
32	1	14	4	181	200	186,5	0	3	4	161	168	163,75
22	4	2	1	185	192	186	1	9	3	207	220	210,75

Tabla 4.8. Resultados de los diez primeros informantes ordenados según el ICoAL en lengua extranjera

Las palabras nucleares (N1) representan la mayor parte con un 92,41 % en lengua española y un 91,11 % en lengua extranjera, cifra ligeramente inferior que muestra que en inglés y francés hay menos asociaciones directas con el núcleo. La distribución porcentual de los otros tres grupos de palabras difieren según el idioma,

²⁰⁸ Vid. apartado 2.5.3 para más detalles.

siendo muy sutil la diferencia: en el N5 en español es de 3,21 % y en lengua extranjera de 3,77 %. Sin embargo, en español aumenta el N25, con un 3,55 % y en lengua extranjera disminuye hasta 2,97 %. Los resultados distan más en el N0 donde en lengua española solo se registra un 0,83 % de las palabras como asociaciones libres frente al 2,16 % en lengua extranjera.

Diez primeros informantes ordenados según ICoAL												
Sujeto	P0	P25	P5	P1	IAL	ICoAL	P0 (LE)	P25 (LE)	P5 (LE)	P1 (LE)	IAL (LE)	ICoAL (LE)
21	1	6	19	268	294	279	2	15	9	186	212	194,25
34	8	3	27	260	298	274,25	26	2	23	145	196	157
16	2	19	28	251	300	269,75	10	14	23	188	235	203
15	6	7	8	252	273	257,75	2	3	4	184	193	186,75
31	2	5	5	254	266	257,75	0	2	3	152	157	154
29	1	7	17	239	264	249,25	1	2	7	146	156	150
8	4	6	21	235	266	247	6	7	13	192	218	200,25
36	3	6	1	245	255	247	2	9	5	142	158	146,75
19	1	15	4	237	257	242,75	12	15	12	162	201	171,75
4	2	14	7	235	258	242	2	6	9	144	161	150

Tabla 4.9. Resultados de los diez primeros informantes ordenados según el ICoAL en lengua española

Cabe destacar que la media del ICoAL en español es bastante superior a la media en lengua extranjera, pasando de un índice de 181,64 en español a 124,89 en lengua extranjera. Asimismo el IAL, o media de las palabras por informante, presenta una media de 191,91 palabras frente a las 133,74 en lengua extranjera²⁰⁹. El valor máximo en lengua española es de 279 palabras frente a 221 en lengua extranjera, por lo que la diferencia de unidades léxicas aportadas en una y otra lengua es sustancial.

Los informantes numerados como 16, 8 y 19 ocupan según el IAL una posición más avanzada que en el ICoAL en lengua española al igual que los sujetos 21 y 32 en lengua extranjera, debido a que han aportado más unidades léxicas que los informantes que le preceden en la lista pero menos vocablos nucleares. De esta forma, se consigue obtener un cálculo que integra la cantidad y la calidad de las palabras aportadas por un individuo.

²⁰⁹ Vid. apartado 4.2.

4.4.2. Comparación intermuestral

4.4.2.1. Estudios en la provincia de Málaga

A continuación, se tendrán en cuenta el promedio de dos estudios realizados en Málaga que pertenecen a los siguientes trabajos:

- Ávila Muñoz (2006), *Léxico disponible de los estudiantes preuniversitarios de Málaga*. Tal y como indica el nombre, el universo está formado por estudiantes preuniversitarios que cursan 2.º de Bachillerato en la provincia de Málaga.
- Ávila Muñoz y Villena Ponsoda (Eds.) (2010), *Variación social del léxico disponible en la ciudad de Málaga*. En este caso, se seleccionaron informantes en la ciudad de Málaga a partir de una muestra de afijación uniforme ($n = 72$) cuyas variables de preestratificación fueron la edad, el sexo y la instrucción.

A pesar de que se trata de estudios con una población y fecha de realización diversas, el rango de algunos centros de interés coinciden. Por ejemplo, el centro más productivo es el tercero, *Comidas y bebidas*; seguido del segundo, *La ropa*. Asimismo, si se obvia el centro relativo a las nuevas tecnologías por no aparecer en el estudio de 2006, también coincidiría en el séptimo puesto el centro número siete, *Juegos y distracciones*; y en el último, el centro número seis, *Medios de transporte*²¹⁰.

	Media 2006	Rango 2006	Media 2010	Rango 2010	Media 2011	Rango 2011
01. Partes del cuerpo	22,4	2	23,04	2	26,28	2
02. La ropa	19,3	5	19,91	5	21,99	3
03. Comidas y bebidas	23,1	1	24,01	1	26,70	1
04. La escuela (muebles y materiales)	20,04	4	19,75	6	21,74	4
05. La ciudad	22,1	3	20,29	4	21,70	5
06. Medios de transporte	15,2	8	17,75	8	15,44	9
07. Juegos y distracciones	16,1	7	18,13	7	17,78	8
08. Profesiones y oficios	19,3	6	21,09	3	21,05	6
09. Ordenadores e internet			17,19	9	19,24	7

Tabla 4.10. Promedio de palabras y rango por centros de interés en español en tres estudios realizados en Málaga (Ávila Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010 y el estudio actual)

En cuanto a los resultados, el promedio del presente estudio es superior en los cuatro primeros centros y en el último, ocupando el segundo lugar en los demás

²¹⁰ La numeración de los centros de interés corresponde a la muestra actual. Las anteriores investigaciones incorporan otros centros de interés además de los coincidentes. En el primer caso, la prueba se realizó con los dieciséis centros de interés tradicionales y, en el segundo, se añadieron los centros *Dinero y economía*, *Internet y ordenadores*, *Los colores* y *El mar*.

centros. Cabe mencionar que las tres investigaciones han seguido las mismas pautas para la realización de las encuestas de disponibilidad léxica en tres momentos diferentes. Por tanto, dado que el nivel de instrucción de los informantes es uno de los factores diferenciadores de los tres estudios, se podría concluir que influye positivamente en la disponibilidad léxica.

Los estudiantes preuniversitarios ocupan el tercer lugar en todos los centros de interés a excepción del cuarto y del quinto, con lo que la muestra general de la provincia de Málaga ocupa el segundo lugar en cuanto a léxico disponible. La media de palabras por informantes del estudio correspondiente al año 2006 es de 19,69 palabras, del año 2010 es de 20,5 y del año 2011, 21,86 palabras. En el recuento se ha omitido el centro nueve porque no se incluyó en el estudio del año 2006.

4.4.2.2. Estudios en lengua extranjera

En este apartado se pretende realizar una comparación del léxico en lengua extranjera con los siguientes estudios:

- Dimitrijević, (1969), *Lexical availability*. Las pruebas las realizaron 185 adolescentes de 14 años, 96 chicos y 89 chicas. El 70 % de los estudiantes cursaban Junior Secondary School y el 30 % restante Senior Secondary School. En este caso, se producen las siguientes diferencias en la traducción realizada por el estudioso yugoeslavo y en la muestra actual: el centro tres es denominado *Food*, el cinco *Town*, el séptimo *Entertainment* y el octavo *Jobs and professions* en lugar de *Food and drink*, *The city*, *Games and entertainment* y *Professions and jobs*. Se ha precisado la diferencia en el último caso por tratarse de un cambio de orden de palabras que puede ver influidos los resultados.
- Ferreira (2006), *Disponibilidad léxica en inglés como lengua materna e inglés como lengua extranjera. Estudio del léxico disponible desde una perspectiva psicolingüística*. En esta investigación de carácter exploratorio, se analizan los datos en inglés de 50 hablantes nativos²¹¹ y 50 alumnos de cuarto curso de las carreras de Pedagogía en Inglés y de Traducción en Idiomas Extranjeros. Existe una ligera diferencia

²¹¹ Estas pruebas fueron realizadas en Inglaterra, concretamente en The Royal School.

en la traducción de dos centros de interés: *Human body parts* y *Games and entertainments* fueron traducidos por Ferreria como *Body parts* y *Entertainment*.

- Carcedo González (2000c), *Disponibilidad léxica en español como lengua extranjera: el caso finlandés*. La muestra está formada por 350 informantes finlandeses cuya lengua materna es el finés o el sueco. El estudioso asturiano divide a los estudiantes según su nivel educativo: 300 escolares y 50 universitarios, de los cuales 25 son de primer ciclo y 25 del segundo. Teniendo en cuenta que en este trabajo los informantes son postgraduados, se han seleccionado los 25 informantes correspondientes al segundo ciclo y que han aprobado más de 30 créditos.
- Samper Hernández (2002), *Disponibilidad léxica en alumnos de español como lengua extranjera*. La investigación está integrada por 45 estudiantes de español como lengua extranjera que estudiaron en Salamanca al menos durante un trimestre. Los centros de interés corresponden a los dieciséis tradicionales.

En la investigación llevada a cabo por Dimitrijević (1969) se confirma la mayor productividad de los informantes. La media registrada en los seis centros coincidentes es de 17,69 en este estudio y 20,04 en el realizado en 1969. No obstante, en los dos primeros casos la media de la muestra actual es superior debiéndose quizás a la mayor edad de los hablantes que les lleva a conocer mejor el tema de la ropa y la comida.

Centros de interés	Estudio actual		Adolescentes (Dimitrijević, 1969)	
	Media	Rango	Media	Rango
02. Clothes	21,99	2	21,81	2
03. Food and drink	26,7	1	28,42	1
05. The city	15,63	4	20,88	3
06. Means of transport	11,17	6	16,98	5
07. Games and entertainments	17,78	3	17,03	4
08. Jobs and professions	12,85	5	15,12	6
Media	17,69		20,04	

Tabla 4.11. Comparación de la media de palabras en inglés del estudio actual y de Dimitrijević (1969)

En el estudio de Ferreira (2006), se constata que los hablantes nativos producen más palabras en todos los centros de interés, obteniendo una media por informante y centro de 27,28 palabras. Le siguen los resultados del estudio actual

cuya media es de 23,19. Por último, se sitúan los estudiantes chilenos con una media de 20,33. En parte, la diferencia entre el promedio de los estudiantes españoles y chilenos puede deberse a que en el primer caso cursan un posgrado mientras que en el segundo caso están estudiando la carrera pero en el primer caso son especialistas en lengua extranjera y en el segundo de diferentes áreas.

Centros de interés	Estudio actual		Nativos		Estudiantes	
	Media	Rango	Media	Rango	Media	Rango
01. Human body parts	26,28	2	26,7	3	21,2	2
02. Clothes	21,99	3	27,8	2	18,8	3
03. Food and drink	26,7	1	29,6	1	23,6	1
07. Games and entertainments	17,78	4	25	4	17,7	4
Media	23,19		27,28		20,33	

Tabla 4.12. Comparación de la media de palabras en inglés del estudio actual y de Ferreira (2006)

En la investigación finlandesa (Carcedo, 2000c) el promedio de palabras por informantes es de 13,68 mientras que en lengua extranjera²¹² los estudiantes en Málaga aportaron de media 17,24. En todos los centros de interés la media en Málaga es superior a excepción del octavo centro, *Profesiones y oficios*, con 13,68 palabras en español frente a 17,24 en lengua extranjera. En cuanto al rango por palabras, no existe ninguna coincidencia en la clasificación.

Centros de interés	Estudio actual		Estudiantes finlandeses	
	Media	Rango	Media	Rango
01. El cuerpo humano	17,53	4	19,4	1
02. La ropa	21,99	2	14,1	4
03. Comidas y bebidas	26,7	1	19	2
04. La escuela	14,29	6	9,2	7
05. La ciudad	15,63	5	14,4	3
06. Medios de transporte	11,17	8	10,8	6
07. Juegos y distracciones	17,78	3	8,7	8
08. Profesiones y oficios	12,85	7	13,8	5
Media	17,24		13,68	

Tabla 4.13. Comparación del promedio de palabras en español y en el estudio de Carcedo (2000c)

Los datos de la investigación de Samper Hernández (2002) muestran que la media aportada por los estudiantes españoles en lengua extranjera es superior al de los estudiantes de español como lengua extranjera. La media general en lengua

²¹² En ambos estudios se analizan las palabras aportadas en lengua extranjera: en español en el caso finlandés y en inglés y francés en el caso de la actual muestra. No obstante, como la redacción es en español se hará referencia a los datos en inglés y francés como lengua extranjera.

extranjera es de 17,24 mientras que en español es de 14,11. Solamente el quinto centro, *La ciudad*, supera la media en lengua española con 17,84 palabras por informante frente a las 15,63 palabras aportadas en lengua extranjera.

Centros de interés	Estudio actual		Extranjeros en Salamanca	
	Media	Rango	Media	Rango
01. El cuerpo humano	17,53	4	16,33	3
02. La ropa	21,99	2	11,66	5
03. Alimentos y bebidas	26,7	1	20,42	1
04. La escuela	14,29	6	11,26	7
05. La ciudad	15,63	5	17,84	2
06. Medios de transporte	11,17	8	11,6	6
07. Juegos y distracciones	17,78	3	11,11	8
08. Profesiones y oficios	12,85	7	12,66	4
Media	17,24		14,11	

Tabla 4.14. Comparación del promedio de palabras en español y en el estudio de Samper Hernández (2002)

En cuanto a la distribución por rangos, ambos estudios coinciden en posicionar el centro *La escuela: muebles y materiales* en primer lugar. A continuación, en lengua extranjera le sigue el centro denominado *La ropa* y en español *La ciudad*, no produciéndose ninguna otra coincidencia. El caso más diverso es el centro *Medios de transporte* ya que en el estudio actual ocupa tanto en español como en lengua extranjera la última posición mientras que en el estudio realizado en Salamanca ocupa el penúltimo lugar.

4.5. Análisis de los vocablos con mayor índice de disponibilidad

Resultaría una tarea encomiable comparar las miles de palabras aportadas en cada encuesta de disponibilidad léxica. Para facilitar el paragón, muchos estudiosos seleccionan los primeros vocablos de la lista ordenada de forma decreciente según el índice de disponibilidad léxica. Siguiendo el patrón de los autores franceses, Gougenheim *et al.* (1967: 165-181), los cuales escogieron los 20 primeros vocablos, otros investigadores como Carcedo González (1999c, 2000a, 2000b, 2000c, 2001); Mateo García (1995) y Bellón Fernández (2011) toman ese mismo criterio.

Sin embargo, en cada estudio el rango puede ser más o menos estricto. Así, Carcedo González además de incluir la comparativa de los 20 primeras lexías, añade una lista con las 10 (Carcedo González, 1999c, 2000b, 2000c) y las 5 palabras más

disponibles (Carcedo González, 1999a, 1999c, 2000c). En el lado opuesto, estarían autores que ampliaron ese rango, como Alba Ovalle (1996), quien trabaja con los primeros 25 vocablos en 1996, con 50 en 1998 y 2000a y con 100 vocablos en 1998 y 2000b. Este último criterio fue empleado por Valencia Espinoza (1998b, 2000), Alba Valencia y Echeverría (1998). En cualquier caso, cuanto menor sea el número de vocablos seleccionados, mayor será el porcentaje de coincidencia y viceversa. Por ello, en las investigaciones de López Chávez (1992, 1995a) cuyas comparaciones comprenden el total de vocablos provenientes de los léxicos de Madrid, Gran Canaria, República Dominicana y Puerto Rico, el porcentaje de coincidencia no alcanza el 25 %.

4.5.1. Comparación intramuestral

Teniendo en cuenta que los mismos informantes han realizado las pruebas de disponibilidad léxica en español y en lengua extranjera, desde el punto de vista de la didáctica de la lengua resulta interesante comparar qué palabras se repiten en español y en lengua extranjera. En las tablas siguientes, las lexías que aparecen en los tres idiomas están resaltadas en el estilo negrita, las que coinciden solo en español e inglés en azul y las que coinciden en francés y español o francés e inglés en verde. La equivalencia entre ellas se puede comprobar al buscarlas en cualquier diccionario bilingüe. Sin embargo, los matices y las connotaciones pueden variar, sobre todo si la búsqueda se realiza por separado en diccionarios monolingües. Por tanto, se ha utilizado la traducción en un sentido amplio, tal y como definen algunos autores:

Le passage d'une langue A à une langue B pour exprimer une même réalité X, passage que l'on dénomme habituellement traduction relève d'une discipline particulière, de nature comparative (Vinay y Darbelnet, 1958: 20).

La traduction consiste à reproduire dans la langue réceptrice le message de la langue source au moyen de l'équivalent le plus proche et le plus naturel, d'abord en ce qui concerne le sens, ensuite en ce qui concerne le style (Taber y Nida, 1971: 11).

Traduire, c'est énoncer dans une autre langue (ou langue cible) ce qui a été énoncé dans une langue source en conservant les équivalences sémantiques et stylistiques (Dubois, Giacomo, Guespin y Marcellesi, 1973: 490).

Traducir es enunciar en otra lengua lo que ha sido enunciado en una lengua fuente, conservando las equivalencias semánticas y estilísticas (García Yebra, 1982: 30).

El primer centro, *Partes del cuerpo humano*, es el que cuenta con un mayor número de palabras coincidentes en los tres idiomas, en concreto 16 de 20. Las palabras que no se repiten son *pecho*, *codo*, *ceja* y *espalda*, aunque esta última se encuentra en el listado francés. Se han destacado en negrita diecisiete palabras en inglés y en francés en el primer centro mientras que en español se han subrayado dieciséis. Ello se debe a que se lematizaron plurales irregulares como vocablos diferentes, es el caso del inglés *foot-feet* y del francés *œil-yeux*.

En el segundo centro, *La ropa*, cuatro de las cinco palabras que no están presentes en los tres listados se refieren a ropa interior: *media*, *braga*, *sujetador* y *calzoncillo*. Las dos primeras coinciden en español con el listado en francés y la tercera con el listado en inglés, por lo que puede deberse a que en español los informantes incluyen las palabras más usadas pero en lengua extranjera las más estudiadas. Así, vocablos como *corbata* (que aparece en inglés y francés) no es una prenda que suelen llevar los estudiantes pero que está presente en los manuales en lengua extranjera.

Centro	Español	Inglés	Francés
01. El cuerpo humano	ojo, cabeza, mano, dedo, pierna, nariz, pie, boca, oreja, brazo, pelo, uña, cuello, hombro, rodilla, pecho, codo, espalda , diente, ceja	head, eye, leg, arm, nose, finger, hand, mouth, ear, hair, shoulder, foot, teeth, neck, nail, feet, toe, knee, lip, heart	tête, main, bras, pied, bouche, oreille, jambe, doigt, nez, œil, yeux, cheveu, cou, ongle, dent, ventre, dos , épaule, genou, langue
02. La ropa	pantalón, camiseta, camisa, falda, jersey, zapato, chaqueta, calcetín, vestido, bufanda, abrigo, media , braga , calzoncillo, guante, sujetador , gorro, (pantalón) vaquero, bota, sombrero	trouser, T-shirt, skirt, shoe, jeans, shirt, sock, jacket, hat, dress, scarf, coat, pullover, boot, glove, shorts, bra , sweater, tie , underwear	pantalon, chaussure, jupe, chemise, t(ee)-shirt, robe, chaussette, veste, manteau, jean, gant, chapeau, pull(-over), écharpe, cravate , culotte , blouse, foulard, collant , botte
03. Comidas y bebidas	agua, tomate , carne, zum , cerveza, pescado, vino, patata, pan, leche, lechuga, arroz, Coca-Cola, café , fruta , naranja, verdura , ensalada, manzana, pollo	water, meat, fish, tomato , potato, wine, milk, Coke, orange, juice , beer, bread, chip, coffee, apple, vegetable , hamburger, banana, chicken , pizza	eau, chocolat, lait, poisson, fromage, pain, pomme, viande, fruit , croissant, omelette, bière, jambon, orange, vin, pomme de terre, crêpe, Coca-Cola, café, confiture
04. La escuela: muebles y materiales	pizarra, silla, mesa, lápiz, boli(grafo), tiza , libro, pupitre , goma (de borrar), ordenador, libreta, regla, borrador, estuche, rotulador , profesor, mochila , cuaderno, alumno , proyector, alumno	pencil, table, chair, pen, book, blackboard, desk , computer, notebook, rubber, teacher, paper , window, chalk , ruler, door, rule, pupil , student, eraser	table, chaise, stylo, tableau, crayon, livre, cahier, professeur, ordinateur, fenêtre , gomme, feutre , règle, papier , porte , dossier, effaceur, feuille, télévision, sac à dos
05. La ciudad	calle, coche, edificio, semáforo, parque, tienda, carretera, casa, plaza, centro comercial, (auto)bus, restaurante, acera, bar,	car, street, building, park, bus, people , shop, house, road, traffic light, square, restaurant, tree, school, museum ,	rue, voiture, bâtiment, route, maison, place, mairie , cinéma, avenue , boulangerie, magasin, vélo, autobus, église ,

	avenida, ayuntamiento, hospital, gente	museo, cine,	supermarket, underground, flat, church	cinema, métro, train, parc	pólice, restaurant, taxi,
06. Medios de transporte	coche, avión, tren, bici(cleta), (auto)bus, moto(cicleta), barco, metro, taxi, tranvía, helicóptero, patín, camión, caballo, furgoneta, yate, monopatín, avioneta, patinete, crucero		car, bus, train, plane, motorbike, bike, boat, taxi, bicycle, ship, underground, helicopter, horse, airplane, on foot, metro, motorcycle, tube, skate, tram	voiture, vélo, avion, train, bus, métro, bateau, tram(way), cheval, moto, TGV, taxi, à pied, RER, TER, camion, tricycle, navette, parachute, minibús	
07. Juegos y distracciones	fútbol, parchís, baloncesto, cine, (juego de) cartas, videojuego, ajedrez, tenis, oca, televisión, leer, deporte, música, teatro, pasear, bailar, internet, juego de mesa, escondite, pádel		football, basketball, cinema, tennis, videogame, music, television, sport, chess, theatre, card, dance, Play Station, reading, swimming, volleyball, running, jogging, baseball, ball	football, basket(-ball), cinéma, tennis, jeu vidéo, théâtre, rugby, volley-ball, Monopoly, natation, sport, Internet, fête, vélo, musique, poker, danser, discothèque, jeu de l'oie, musée	
08. Profesiones y oficios	profesor, enfermero, policía, maestro, arquitecto, fontanero, panadero, bombero, pintor, albañil, conductor, electricista, taxista, carpintero, actor	médico, abogado, ingeniero, camarero, cocinero, panadero, conductor, electricista, taxista, carpintero, actor	teacher, doctor, policeman, nurse, lawyer, engineer, driver, waiter, architect, fireman, actor, professor, plumber, writer, painter, taxi driver, singer, butcher, shop assistant, cook	professeur, boulangier, policier, pompier, cuisinier, médecin, boucher, plombier, docteur, chef, infirmier, vendeur, architecte, traducteur, pâtissier, chauffeur, dentiste, maître, serveur, acteur	
09. Ordenadores e internet	ratón, teclado, pantalla, página web, portátil, correo (electrónico), disco duro, torre, impresora, Facebook, red social, monitor, programa, cable, archivo, software, pendrive, hardware, Tuenti, Google		mouse, screen, keyboard, Facebook, laptop, web, email, hardware, software, Google, download, chat, computer, website, program(me), Tuenti, Twitter, pendrive, Internet, printer	souris, clavier, écran, Internet, Facebook, page web, Google, logiciel, clé USB, ordinateur, chat, câble, site (web), imprimante, JClick, CD, chatter, Hot Potatoes, réseau social, Wi-Fi	

Tabla 4.15. Comparación de los 20 vocablos más disponibles en español, inglés y francés

En el tercer centro, *Comidas y bebidas*, las respuestas se diversifican. En los tres listados coinciden 12 palabras, entre los que destaca el nombre de la marca de bebidas, *Coca-Cola*. Además, en inglés se incorporan cuatro palabras que aparecen en español (*tomato, juice, vegetable* y *chicken*) y en francés solo una (*fruit*). Por su parte, en francés se ha incluido la palabra *crêpe* que es un extranjerismo ya acuñado en español pero no es tan disponible para los informantes en su lengua materna.

En el cuarto centro, *Escuela: muebles y materiales*, se han marcado trece coincidencias en español pero solo doce en lengua extranjera. En este caso, las palabras *libreta* y *cuaderno* se refieren en los idiomas extranjeros a una misma realidad: en inglés a *notebook*, y en francés a *cahier*. Además, en lengua extranjera se han incluido palabras genéricas como *ventana* y *puerta* que no aparecen en el listado en español,

quizás por carecer de un vocabulario disponible específicamente relacionado con este centro.

El quinto centro, *La ciudad*, es un centro poco cohesionado pero mantiene la proporción del 60 % de palabras coincidentes en los tres listados. En el centro seis, *Medios de transporte*, se han subrayado varios términos con sinónimos en inglés: *motorcycle* y *motorbike*, que se refieren a *moto(cicleta)* en español y a *moto(cyclette)* en francés; *bike* y *bicycle* que se refieren a *bici(cicleta)* en español y *vélo*²¹³ en francés; *tube*, *underground* y *metro* que se refieren en español a metro y en francés a *métro*; *plane* y *airplane* las cuales se refieren a *avión* en español y *avion* en francés aunque *airplane* también podría referirse a *aeroplano*. Por ello, están marcadas dieciséis palabras en negrita en inglés mientras que en español y francés solo once.

En el séptimo centro, *Juegos y distracciones*, tan solo coinciden 9 palabras en los tres listados. En inglés aparecen como disponibles palabras que se han incorporado al repertorio lingüístico del español, como *running*, *jogging*, *tennis*, *volleyball*, *football*, *basketball* y *baseball* además de un nombre propio, *Play Station*. En francés, se incluye también el nombre de un juego, *Monopoly*, además de préstamos del inglés: *footbal*, *basket(-ball)*, *volley-ball*, *tennis*, *cinéma*, *rugby* y *poker*.

En el centro octavo, *Profesiones y oficios*, el vocablo más disponible llama la atención porque en los tres listados es *profesor*, profesión para la que se prepara el alumnado cursa el máster. En inglés, los informantes han escrito *teacher* y *professor* pero en español equivale a *profesor* aunque con el matiz que *professor* se refiere a profesor universitario. En francés aparecen doce palabras marcadas en el centro ocho, ya que los vocablos *docteur* y *médecin* se refieren en inglés a doctor. En español se usa *doctor* y *médico* respectivamente.

En el noveno centro, *Ordenadores e internet*, forman parte de los tres listados las palabras *Facebook* y *Google*, de ahí que se intuya el uso que hacen los estudiantes de las nuevas tecnologías: para comunicarse con amigos y buscar información. En español y en francés se han incorporado una parte considerable de préstamos del inglés

²¹³ También aparece en francés el vocablo *bicyclette*, pero no se sitúa entre los veinte primeros por ser un arcaísmo.

(*hardware, software*, etc.) y, además, en francés se han anotado nombres de herramientas informáticas utilizadas con fines pedagógicos: *Hot potatoes* y *JClic*.

CI	ES/IN/FR		ES/IN		ES/FR		IN/FR	
	Vocablos	%	Vocablos	%	Vocablos	%	Vocablos	%
1	16	80 %	0	0 %	1	5 %	0	0 %
2	15	75 %	1	5 %	2	10 %	1	5 %
3	12	60 %	4	20 %	1	5 %	0	0 %
4	12	60 %	3	15 %	2	10 %	3	15 %
5	12	60 %	2	10 %	2	10 %	2	10 %
6	11	55 %	2	10 %	1	5 %	1	5 %
7	9	45 %	4	20 %	2	10 %	2	10 %
8	12	60 %	4	20 %	1	5 %	2	10 %
9	10	50 %	4	20 %	2	10 %	2	10 %
Media	12,22	61,11 %	2,67	13,33 %	1,56	7,78 %	1,4	7,22 %

Tabla 4.16. Vocablos coincidentes entre los 20 más disponibles en español, francés e inglés

En síntesis, se confirma que más de la mitad de las palabras coinciden en los tres idiomas con un 61,11 % de media. En cuanto a las lexías marcadas exclusivamente en dos idiomas, se observa que la combinación español e inglés presenta el mayor porcentaje con un 13,33 %. Los resultados en los otros pares de idiomas son muy similares ya que en español y francés es de 7,78 % y en inglés y francés es ligeramente inferior, 7,22 %. A través del análisis del índice de cohesión por centros, se constata que el primero muestra datos cohesionados en los tres idiomas mientras que el centro siete presenta un léxico más diverso.

4.5.2. Comparación intermuestral

4.5.2.1. Investigaciones en español

a) Comparación con estudios con estudiantes extranjeros

En este apartado se analizan las palabras más disponibles aportadas en este estudio y en las investigaciones llevadas a cabo por Carcedo González (2000c) y Samper Hernández (2002). El factor que marca la diferencia entre las tres muestras es la forma de adquisición y aprendizaje de la lengua: en Carcedo González (2000c) se trata de estudiantes finlandeses de español como lengua extranjera y en Samper Hernández (2002) de estudiantes extranjeros de español de diferentes nacionalidades

que realizan una estancia en España y en este estudio por hablantes nativos. Del mismo modo, se pueden establecer diferencias entre el grado de exposición a la lengua en estudiantes, ya que en el caso de Samper Hernández (2002) los informantes necesitan utilizar el español como lengua vehicular en situaciones cotidianas donde se requiere el uso del vocabulario disponible.

Centro	Carcedo (2000c): estudiantes finlandeses	Samper (2002): estudiantes extranjeros en España	Estudio actual: estudiantes de postgrado
01. El cuerpo humano	mano, cabeza, ojo, pie, pelo, boca, nariz, dedo, estómago , espalda, oreja, labio , pierna, corazón, cara , brazo, rodilla, diente, uña, frente	ojo, cabeza, pie, pierna, mano, nariz, brazo, dedo, oreja, boca, pelo, cara , estómago , diente, labio , espalda, pecho , rodilla, cuello , uña	ojo, cabeza, mano, dedo, pierna, nariz, pie, boca, oreja, brazo, pelo, uña, cuello , hombro, rodilla, pecho , codo, espalda, diente, ceja
02. La ropa	zapato, pantalón, vaquero, camisa, blusa, chaqueta, camiseta, vestido, falda, jersey, calcetín, sombrero, traje , traje de baño, gorro , corbata , zapatilla , abrigo, bota , bufanda	pantalón, camiseta, camisa, zapato, chaqueta, falda, jersey, abrigo, calcetín, vestido, bufanda, corbata , vaquero, sombrero, guante , suéter, traje , zapatilla , ropa interior, braga	pantalón, camiseta, camisa, falda, jersey, zapato, chaqueta, calcetín, vestido, bufanda, abrigo, media, braga , calzoncillo, guante , sujetador, gorro , (pantalón) vaquero, bota , sombrero
03. Comidas y bebidas	leche, agua, cerveza, carne, vino, café, pan, pescado, patata, manzana, paella, naranja, té , queso , zumo, tortilla, huevo, (vino) tinto, pollo, uva	agua, zumo, carne, cerveza, pan, leche, pollo, pescado, café, Coca-Cola , tomate , vino, naranja, manzana, pasta, queso , patata, verdura , té , jamón	agua, tomate , carne, zumo, cerveza, pescado, vino, patata, pan, leche, lechuga, arroz, Coca-Cola , café, fruta, naranja, verdura , ensalada, manzana, pollo
04. La escuela: muebles y materiales	libro, lápiz, papel , bolí(grafo), mesa, goma, silla, cuaderno, bolsa, mochila, bolso, diccionario , pizarra, mapa, pluma , tele(visión) , ordenador, transparencia, borrador, sacapuntas	mesa, silla, bolí(grafo), lápiz, libro, cuaderno, papel , pizarra, tiza , tele(visión) , diccionario , luz, borrador, ventana, mochila, ordenador, pluma , goma (de borrar), puerta, escritorio	pizarra, silla, mesa, lápiz, bolí(grafo), tiza , libro, pupitre, goma (de borrar), ordenador, libreta, regla, borrador, estuche, rotulador, profesor, cuaderno , mochila, proyector, alumno
05. La ciudad	calle, coche, casa, tienda, escuela , parque, mercado, iglesia , supermercado , gente, restaurante, banco, autobús, plaza, cine , tráfico, edificio, correo, hospital , museo	calle, coche, tienda, edificio, plaza, iglesia , parque, casa, gente, (auto)bús, restaurante, bar , semáforo , avenida , escuela , super(mercado) , universidad, discoteca, árbol, farmacia	calle, coche, edificio, semáforo , parque, tienda, carretera, casa, plaza, centro comercial , (auto)bus, restaurante, acera, bar , avenida , museo , ayuntamiento, cine , hospital , gente
06. Medios de transporte	coche, autobús, tren, avión, bici(cleta), barco, metro, taxi, tranvía , a pie, barca, moto, camión, motocicleta, caballo, carro , aeroplano, auto, helicóptero , tractor	coche, (auto)bús, tren, avión, bici(cleta), barco, taxi, metro, pie , moto, camión, motocicleta, caballo, autostop, andar, autocar, correr, aeropuerto, tráfico, carro	coche, avión, tren, bici(cleta), (auto)bus, moto(cicleta), barco, metro, taxi, tranvía , helicóptero , patín, camión, caballo, furgoneta, yate, monopatín, avioneta, patinete, crucero
07. Juegos y distracciones	fútbol, tenis, tele(visión) , cine, bailar, teatro , música, nadar , jugar, leer, voleibol, deporte, ajedrez	fútbol, tenis, baloncesto, (jugar a las) cartas, leer , bailar, nadar , béisbol, correr, esquiar, golf , deporte, rugby,	fútbol, parchís, baloncesto, cine, (juego de) cartas, videojuego, ajedrez , tenis, oca, televisión , leer ,

	libro, guitarra, baloncesto , película, cantar, cartas , golf	natación, hockey, ver televisión ²¹⁴ , escuchar música, Monopoly, (ir al) cine, televisión	deporte , música , teatro , pasear, bailar , internet, juego de mesa, escondite, pádel
08. Profesiones y oficios	profesor , médico , policía , vendedor, camarero , dentista , jardinero, enfermero , taxista , secretario , doctor , actor , bailarín, abogado , cura, pintor, cocinero, bombero, peluquero , jefe	profesor , médico , abogado , estudiante, camarero , policía , maestro, ama de casa, secretario , enfermero , peluquero , ingeniero , cantante, escritor, doctor , actor , piloto, director, zapatero, cocinero	profesor , médico , enfermero , abogado , policía , ingeniero , maestro, camarero , arquitecto, cocinero , fontanero, panadero, bombero, pintor, albañil, conductor, electricista, taxista , carpintero, actor

Tabla 4.17. Comparación de los veinte primeros vocablos en español de las investigaciones de Carcedo González (2000c), Samper Hernández (2002) y del estudio actual

Algunas de las peculiaridades comentadas en el apartado anterior se repiten. Por ejemplo, entre los vocablos que no coinciden en los tres listados, se destaca en el centro *La ropa*, palabras relacionadas con la ropa interior (*media*, *braga*, *calzoncillo* y *sujetador*) mientras que *corbata* solo aparece en los listados de español como lengua extranjera. Asimismo, la profesión más disponible es *profesor*, quizás por encontrarse en un ámbito académico. En el centro referido a los medios de transporte, se localiza la palabra *carro* en los dos estudios de español como lengua extranjera, Carcedo González (2000c) y Samper Hernández (2002) pero no así en el listado de hablantes nativos.

CI	Tres estudios		2000/2002		2000/2011		2002/2011	
	Vocablos	%	Vocablos	%	Vocablos	%	Vocablos	%
1	15	75 %	3	15 %	0	0 %	2	10 %
2	13	65 %	3	15 %	2	10 %	2	10 %
3	13	65 %	2	10 %	0	0 %	2	10 %
4	11	55 %	4	20 %	0	0 %	1	5 %
5	10	50 %	3	15 %	3	15 %	3	15 %
6	12	60 %	2	10 %	0	0 %	2	10 %
7	9	45 %	2	10 %	1	5 %	2	10 %
8	8	40 %	3	15 %	1	5 %	2	10 %
Media	11,38	56,88 %	2,75	13,75 %	0,88	4,38 %	2	10 %

Tabla 4.18. Datos comparativos de las investigaciones de Carcedo González (2000c), Samper Hernández (2002) y del estudio actual

El porcentaje de coincidencia entre las veinte palabras más disponibles de los tres estudios desciende con respecto a la comparación entre los listados de este

²¹⁴ En este estudio se ha lematizado de forma diferente «televisión» y «ver la televisión».

estudio a 56,88 %. Según los datos del índice de cohesión, el primer centro es el más compacto y, por tanto, es el que presenta mayor número de palabras coincidentes en los diferentes listados. A continuación, le siguen el segundo y el tercero, *La ropa* y *Comidas y bebidas*, con más 13 vocablos. El centro más diverso es el octavo, *Profesiones y oficios*, con tan solo 8 vocablos coincidentes

En cuanto a la similitud por pares de investigaciones, las más cercanas son las dos realizadas a estudiantes de español como lengua extranjera. Al 56,88 % de vocablos coincidentes en los tres idiomas, le añaden un 13,75 %. Sin embargo, el porcentaje entre el estudio actual y el estudio realizado en Salamanca es del 10 % y entre el estudio actual y el realizado en Finlandia desciende a 4,38 %.

b) Investigaciones en la provincia de Málaga

Antes de presentar la tabla con los vocablos más disponibles, cabe destacar la diferencia de nomenclatura de las investigaciones del año 2006 y 2010 con respecto al estudio actual. El primer centro de interés se llama *El cuerpo humano* en lugar de *Partes del cuerpo humano*; el quinto *Alimentos* mientras que en esta investigación se denomina *Comidas y bebidas*. En cuanto al último centro, solo se ha podido realizar la comparación entre el año 2010 y 2011, por lo que los datos cuantitativos correspondientes a las otras combinaciones se han realizado en base a ocho centros de interés.

Centro	Preuniversitarios	Población estratificada	Estudiantes de postgrado
01. El cuerpo humano	cabeza, pierna, brazo, mano, ojo, pie, dedo, nariz, boca, oreja, rodilla , uña, pelo, cuello , tronco, codo , hombro , espalda , diente , muñeca	ojo, brazo, pierna, mano, cabeza, pie, nariz, dedo, corazón, boca, uña, oreja, pelo, hígado, pulmón, estómago, hueso, riñón, músculo, ceja	ojo, cabeza, mano, dedo, pierna, nariz, pie, boca, oreja, brazo, pelo, uña, cuello , hombro , rodilla , pecho, codo , espalda , diente , ceja
02. La ropa	pantalón, camiseta, camisa, jersey, calcetín, falda, chaqueta, zapato, braga, abrigo, calzoncillo, sujetador, bufanda, vestido, sudadera, chándal, guante , media, (pantalón) vaquero, blusa	pantalón, camisa, camiseta, calcetín, falda, jersey, abrigo, zapato, chaqueta, braga, sujetador, calzoncillo, blusa , vestido, bufanda, media, corbata, bañador, zapatilla, (pantalón) vaquero	pantalón, camiseta, camisa, falda, jersey, zapato, chaqueta, calcetín, vestido, bufanda, abrigo, media, braga, calzoncillo, guante , sujetador, gorro, (pantalón) vaquero, bota, sombrero
03. Comidas y bebidas	agua , Coca-Cola , carne, zum , pan, pescado, patata, Fanta, tomate, lenteja , leche, lechuga, cerveza , güisqui, vino , arroz, refresco, fruta, huevo, macarrón	pan, leche, tomate, carne, lenteja , patata, pescado, arroz, garbanzo, huevo, lechuga, pimiento, verdura , cebolla, manzana , naranja , fruta, aceite, plátano, queso	agua , tomate, carne, zum , cerveza , pescado, vino, patata, pan, leche, lechuga, arroz, Coca-Cola , café, fruta, naranja , verdura , ensalada, manzana , pollo

04. La escuela: muebles y materiales	silla, pizarra, mesa, boli(grafo), lápiz, libro, goma (de borrar), tiza, pupitre, libreta, estuche , borrador (de pizarra), folio , mochila , carpeta , ventana, sacapuntas , armario, puerta, regla	pizarra, lápiz, silla, mesa, libro, tiza, pupitre, boli(grafo), goma (de borrar), libreta, cuaderno , regla, sacapuntas , folio , borrador (de tiza), pluma, ordenador , lápiz de color, rotulador , carpeta	pizarra, silla, mesa, lápiz, boli(grafo), tiza, libro, pupitre, goma (de borrar), ordenador , libreta, regla, borrador, estuche , rotulador , profesor, cuaderno , mochila , proyector, alumno
05. La ciudad	coche, tienda, calle, edificio, parque, moto(cicleta) , carretera, casa, semáforo, piso, gente, centro comercial , contaminación, (auto)bus, acera , colegio, bar , ruido, restaurante , cine	calle, coche, edificio, parque, semáforo, plaza , tienda, avenida , farola, (auto)bus, jardín, carretera, casa, moto(cicleta) , árbol, banco, bici(cleta), gente, iglesia	calle, coche, edificio, semáforo, parque, tienda, carretera, casa, plaza , centro comercial , (auto)bus, restaurante , acera , bar , avenida , museo, ayuntamiento, cine , hospital, gente
06. Medios de transporte	coche, moto(cicleta) , auto(bus), avión, bici(cleta), tren, barco, camión, taxi, patín, metro, helicóptero, monopatín, furgoneta, patinete, avioneta, ciclomotor, tranvía, triciclo, barca	coche, avión, moto(cicleta) , auto(bus), bici(cleta), tren, barco, patín, helicóptero, camión, metro, avioneta, tranvía, caballo , patinete, carro, furgoneta, globo, taxi, monopatín	coche, avión, tren, bici(cleta), (auto)bus, moto(cicleta) ²¹⁵ , barco, metro, taxi, tranvía, helicóptero, patín, camión, caballo , furgoneta, yate, monopatín, avioneta, patinete, crucero
07. Juegos y distracciones	fútbol, parchís, baloncesto, carta, tele(visión), ordenador, escondite, cine, ajedrez, pilla-pilla , oca, tenis, comba , música , leer, videojuego , volei(bol), videoconsola, internet , Monopoly	parchís, fútbol, oca, ajedrez, cartas, baloncesto, escondite, cine, dominó, tele(visión), tenis, pilla-pilla , pasear , mus, leer, póquer, damas, balonmano, comba , teatro	fútbol, parchís, baloncesto, cine, (juego de) cartas, videojuego , ajedrez, tenis, oca, televisión, leer, deporte, música , teatro , pasear , bailar, internet , juego de mesa, escondite, pádel
08. Profesiones y oficios	profesor/ra, médico/ca, albañil, abogado/da, camarero/ra, policía, maestro/tra, carpintero, fontanero/ra, electricista, amo/ma de casa, enfermero/ra, empresario/ria, agricultor/ra, bombero , jardinero/ra , secretario/ria, cocinero/ra, ingeniero, dependiente/ta	médico/a, albañil, profesor/ora, fontanero, abogado/a, carpintero/a, electricista, maestro/a, enfermero/a, ingeniero, mecánico, pintor , camarero/a, dependiente/a , policía, jardinero , conductor/ora , bombero , cocinero/a, arquitecto/a	profesor, médico, enfermero, abogado, policía, ingeniero, maestro, camarero, arquitecto , cocinero, fontanero, panadero, bombero , pintor , albañil, conductor , electricista, taxista, carpintero, actor
09. Ordenadores e internet		ratón , pantalla , teclado , torre , impresora , página web , chat, ordenador, Messenger, chatear, correo (electrónico) , disco duro , programa , Google , monitor , altavoz, web, software , Internet, CD-ROM	ratón , teclado , pantalla , página web , portátil, correo (electrónico) , disco duro , torre , impresora , Facebook, red social, monitor , programa , cable, archivo, software , pendrive, hardware, Tuenti, Google

Tabla 4.19. Comparación de los veinte primeros vocablos en español de las investigaciones de Ávila Muñoz (2006), Ávila Muñoz y Villena Ponsoda (2010) y del estudio actual

²¹⁵ En la investigación actual se ha lematizado *moto(cicleta)* como un único vocablo mientras que en el resto se recogen por separado *moto* y *motocicleta*.

En el primer centro, *Partes del cuerpo humano*, los listados que guardan mayor relación son el primero (Ávila Muñoz, 2006) y el tercero (este estudio). Los informantes del estudio de Ávila Muñoz y Villena Ponsoda (2010) han aportado más palabras relacionadas con los órganos y partes internas quizás porque la media de edad es mayor y, por tanto, la exposición a diferentes enfermedades también lo es: *corazón, hígado, pulmón, estómago, hueso, pulmón y riñón*.

Como los tres listados corresponden al español como lengua materna, en el segundo centro, *La ropa*, se incluyen las palabras referidas a la ropa interior que no se encontraban en los listados de lengua extranjera: *media, braga, calzoncillo y sujetador*. En cambio, *corbata* solo aparece en el listado del estudio de 2010 ya que hay informantes que han comenzado su vida laboral y que son mayores que en el resto de grupos. En esa misma línea, cabe mencionar que en el centro referido a la comida, aparecen los vocablos *vino, cerveza y Coca-Cola* solo en los listados de preuniversitarios y de estudiantes de postgrado, quizás por ser más frecuentes este tipo de bebidas en los jóvenes.

En el sexto centro, *Medios de transporte*, no aparece la palabra *carro* en ninguno de los listados. Se confirma la presencia en manuales de español de variedades del español de América que no son tan frecuentes en el español peninsular, como *carro* en lugar de *coche*. En cuanto a las profesiones, se repite en primer lugar *profesor* en el estudio con estudiantes (preuniversitarios y de postgrado). En la investigación general de 2010, esta profesión ocupa el tercer lugar ya que son informantes que no están necesariamente vinculados al ámbito académico.

El porcentaje de coincidencia en las tres investigaciones asciende al 63,13 % entre los 20 primeros vocablos según el índice de disponibilidad léxica. Esa cifra aumenta en los centros más compactos, como es el caso del sexto centro, denominado *Medios de transporte* con un 85 %. Sin embargo, el centro más disperso, *Alimentos*, solo cuenta con un 45 % de palabras comunes.

CI	Tres estudios		2006/2010		2006/2011		2010/2011	
	Vocablos	%	Vocablos	%	Vocablos	%	Vocablos	%
1	12	60 %	0	0 %	6	30 %	1	5 %
2	16	80 %	1	5 %	1	5 %	0	0 %
3	9	45 %	1	5 %	5	25 %	3	15 %
4	12	60 %	3	15 %	2	10 %	3	15 %

5	11	55 %	1	5 %	4	20 %	2	10 %
6	17	85 %	0	0 %	0	0 %	1	5 %
7	11	55 %	2	10 %	3	15 %	2	10 %
8	13	65 %	2	10 %	1	5 %	4	20 %
9	0	0 %	0	0 %	0	0 %	12	60 %
Media	12,63	63,13 %	1,25	6,25 %	2,75	13,75 %	3,11	15,56 %

Tabla 4.20. Datos comparativos de las investigaciones de Ávila Muñoz (2006), Ávila Muñoz y Villena Ponsoda (2010) y del estudio actual

Las investigaciones que presentan vocablos más similares son las dos últimas con un 15,56 %, siendo a su vez las que guardan menor diferencia temporal. Sin embargo, los datos de 2006 y 2010 son más diversos, con un 6,25 % de coincidencia, que los de 2006 y 2010. La mayor diferencia entre las tres investigaciones reside en la edad y el nivel de estudios de los informantes: en la primera se trata de estudiantes preuniversitarios, en la segunda de habitantes de diversas edades y niveles de estudio y en la tercera estudiantes de postgrado.

4.5.2.2. Investigaciones en inglés y francés

En lengua extranjera se toman como referencia los estudios de Ferreira (2006) y Dimitrijević (1969) para el inglés y para el francés el trabajo pionero de Gougenheim *et al.* (1964). Al igual que en lengua española, se parte de la hipótesis que el vocabulario de estudiantes de lengua extranjera inglés de dos muestras diferentes, Ferreira (2006) y el estudio actual) guardarán mayor relación que las muestras con estudiantes nativos de Ferreira (2006) y Dimitrijević (1969). En primer lugar, se analiza el léxico más disponible del estudio de Dimitrijević (1969) sobre una muestra de 185 adolescentes de 14 años (96 chicas y 89 chicos) cuyos listados se presentan según sexo.

Centro	Estudio actual	Dimitrijević (1969): Boys	Dimitrijević (1969): Girls
02. Clothes	trouser, T-shirt, skirt, shoe, jeans , shirt, sock, jacket, hat, dress, scarf, coat, pullover , boot, glove, shorts , bra, sweater, tie, underwear	shirt, tie, trousers, socks, shoes, vest, hat, skirt, jacket, scarf, jersey, suit , blouse , pullover , coat, cardigan , jeans , nylons , dress, gloves	blouse , cardigan , skirt, dress, tie, socks, trousers, scarf, jacket, jumper, shirt, hat, shoes, gloves, coat, underskirt, nylons , suit , shorts , blazer
03. Food and drink	water , meat, fish, tomato, potato, wine, milk, Coke, orange, juice, beer, bread, chip, coffee , apple, vegetable, hamburger, banana ,	potatoes, carrots , milk, apples, cabbage , oranges, water , bread, eggs , meat, lettuce , fish, tomatoes, onions , pears, tea , bananas , mutton, peas , leeks	potatoes, carrots , apples, milk, turnip, oranges, bread, eggs , peas , fish, meat, cabbage , beans, tomatoes, butter, tea , cakes, lettuce , onions , coffee

	chicken, pizza		
05. The city	car, street, building, park, bus, people, shop, house, road, traffic light, square, restaurant, tree, school, museum, supermarket, cinema, underground, flat, church	shop, bus, school, house, people, cars, street, road, factory, policeman, smoke, traffic lights, park, building, church, cinema, bicycle, chimney, pavement, lamp-post	shop, bus, people, cars, house, school, road, factory, garden, church, traffic lights, policeman, office, park, street, zebra crossing, building, hospital, cinema, smoke
06. Means of transport	car, bus, train, plane, motorbike, bike, boat, taxi, bicycle, ship, underground, helicopter, horse, airplane ²¹⁶ , on foot, metro, motorcycle, tube, skate, tram	car, bus, trains, aeroplanes, bicycle, ship, horses, lorry, scooter, planes, taxi, walking, boat, punt, sledge, helicopter, motor bike, feet, hovercraft, liner	car, bus, train, horses, aeroplanes, lorry, ship, bicycle, motor bike, scooter, pram, trams, boat, taxi, helicopter, van, donkey, planes, walking, camels
07. Games and entertainment	football, basketball, cinema, tennis, videogame, music, television, sport, chess, theatre, card, dance, Play Station, reading, swimming, volleyball, running, jogging, baseball, ball	TV, football, dancing, theatre, cinema, show, swimming, rugby, singers, snooker, film, play, bingo, records, actors, boxing, radio, singing, wrestling, clubs	dancing, TV, theatre, bassoon, play, show, skating, swimming, records, circus, cinema, film, singing, singers, actors, tennis, hockey, carnival, football, comedian
08. Professions and Jobs	teacher, doctor, policeman, nurse, lawyer, engineer, driver, waiter, architect, fireman, actor, professor, plumber, writer, painter, taxi driver, singer, butcher, shop assistant, cook	teacher, plumber, joiner, policeman, electrician, bus driver, engineer, bus conductor, mechanic, doctor, butcher, sailor, printer, clerk, baker, shop assistant, pilot, postman, milkman, ambulance man	nurse, typist, doctor, teacher, shop assistant, policeman, engineer, printer, electrician, secretary, clerk, bus conductor, plumber, shopkeeper, dentist, joiner, bus driver, telephonist, ambulance man, receptionist

Tabla 4.21. Comparación de los veinte primeros vocablos en inglés de las investigaciones de Dimitrijević (1969) y del estudio actual

El nivel de coincidencia de los vocablos es menor que en las comparaciones anteriores. El porcentaje en los tres listados es del 37,5 % pero los resultados por pares tan solo coinciden con el español en un 2,5 % y un 3,33 %. Esa diferencia se explica por la incidencia en el léxico disponible de tres variables principalmente: la edad de la población objeto de estudio, adolescentes de 14 años y estudiantes españoles de más de 22 años; la adquisición del idioma, lengua materna o lengua extranjera; el grado de instrucción, estudiantes de secundaria y estudiantes postgraduados. Además, se podría tener en cuenta la sincronía de la lengua ya que la primera prueba se realizó en la década de los 60 y la segunda en el año 2011.

En el centro referido a las profesiones, *profesor* aparece en primer lugar en el listado de los chicos y del estudio actual mientras que las chicas han escrito *nurse*

²¹⁶ En Reino Unido se escribe también *aeroplane*, por lo que he subrayado *aeroplane* como si fuese la misma palabra a pesar de la diferente grafía.

seguida por *typist* por tratarse de profesiones relacionadas con el sexo femenino. En cambio, los listados entre chicos y chicas presentan una mayor coincidencia, concretamente un 10 % ya que existen ciertas palabras relacionadas con los jóvenes, como *scooter*, y otras que aparecen en los tres listados pero en posiciones más adelantadas en el caso de los jóvenes ingleses, como *bus* y *school* en el centro denominado *The city*.

CI	Tres listados		Actual/boys		Actual/girls		Boys/girls	
	Vocablo	%	Vocablo	%	Vocablo	%	Vocablo	%
2	12	60 %	2	10 %	1	5 %	5	25 %
3	8	40 %	2	10 %	1	5 %	7	35 %
4	13	65 %	0	0 %	0	0 %	3	15 %
5	12	60 %	0	0 %	1	5 %	3	15 %
6	6	30 %	0	0 %	1	5 %	6	30 %
7	6	30 %	1	5 %	1	5 %	6	30 %
Media	9,50	37,5 %	0,83	2,5 %	0,83	3,33 %	5	10 %

Tabla 4.22. Datos comparativos de la investigación de Dimitrijević (1969) y del estudio actual

A continuación se presenta la comparación con los centros coincidentes de la investigación llevada a cabo por Ferreira (2006). La segunda columna muestra los resultados extraídos de hablantes nativos y la tercera de estudiantes universitarios chilenos.

Centro	Estudio actual	Ferreira (2006): hablantes nativos	Ferreira (2006): estudiantes chilenos
01. Human body parts	head, eye, leg, arm, nose, finger, hand, mouth, ear, hair, shoulder, foot , teeth , neck, nail , feet, toe , knee, lip , heart	leg, arm, head, eye, finger, hand, nose, toe , ear, mouth, feet, knee, heart , neck, shoulder, lung, hair, stomach, ankle, elbow	head, arm, leg, hand, eye, nose, mouth, finger, ear, hair, nail , neck, foot , teeth , tongue, knee, feet, lip , shoulder, face
02. Clothes	trouser, T-shirt, skirt, shoe, jeans, shirt, sock, jacket, hat, dress , scarf , coat , pullover, boot , glove , shorts, bra , sweater , tie , underwear	skirt, shoe, trousers, jeans, t-shirt, shirt, top, fashion, sock, shorts, hat, tights, shopping, money, bra , topshop, shop, colour pants, designer, jacket	pants, shoe, t-shirt, jacket, trousers, sweater , scarf , skirt, jeans, sock, shirt, hat, coat , underwear , dress , blouse, glove , boot , shorts, tie
03. Food and drink	water, meat , fish , tomato , potato , wine, milk, Coke , orange, juice , beer , bread, chip , coffee , apple, vegetable, hamburger, banana , chicken, pizza	water, apple, orange, coke , chip , chocolate, pasta, banana , carrot, wine, chicken, fruit , bread, orange-juice, fat, lemonade, milk, pork, rice , vegetable	meat , tomato , apple, potato , bread, milk, chicken, lettuce, juice , orange, coffee , beer , vegetable, fish , tea, water, wine, soda, rice , fruit
07. Games and entertainment	football, basketball, cinema, tennis, videogame, music, television, sport, chess, theatre, card, dance , Play Station, reading, swimming, volleyball, running, jogging, baseball, ball	tv, film, radio , theatre, music, cinema, magazine, media, movie , fun, show , television, friend , comedy, play , book , party , money, sport, newspaper	cinema, movie , music, theatre, tv , friend , party , television, game, fun, radio , book , sport, concert, park, dancing , play , show , disco, shopping

Tabla 4.23. Comparación de los veinte primeros vocablos en inglés de las investigación de Ferreira (2006) y del estudio actual

En este caso, la coincidencia de palabras en las tres investigaciones no llega al 50 %. Los resultados más similares son los correspondientes a estudiantes chilenos y españoles con un 25 % de compatibilidad aparte del 47,5 % común a los tres listados. Sin embargo, el porcentaje de coincidencia con estudiantes nativos desciende a la mitad: 12,5 % de los vocablos aportados por estudiantes chilenos y a 7,5 % por los españoles. El hecho de que la nomenclatura fuese diversa en el centro referido a las distracciones en español se hace patente en los resultados. El número de coincidencias entre los hablantes nativos y los estudiantes chilenos es del 40 % mientras que entre los hablantes nativos y los estudiantes españoles del 0 % y solo de un 5 % entre los estudiantes chilenos y españoles.

CI	Tres investigaciones		Estudio actual y Hablantes nativos		Estudio actual y estudiantes chilenos		Hablantes nativos y estudiantes chilenos	
	Vocablo	%	Vocablo	%	Vocablo	%	Vocablo	%
1	14	70 %	2	10%	4	20 %	0	0 %
2	10	50 %	1	5 %	8	40 %	0	0 %
3	9	45 %	3	15 %	7	35 %	2	10 %
7	5	25 %	0	0 %	1	5 %	8	40 %
Media	9,5	47,50 %	1,5	7,50 %	5	25 %	2,5	12,50 %

Tabla 4.24. Datos comparativos de la investigación de Ferreira (2006) y del estudio actual

En el estudio de Gougenheim *et al.* (1964) solo se comparan los veinte primeros vocablos por no contar con el total de palabras por centros de interés. Los datos fueron recogidos en diferentes zonas de Francia. En el Département de la Dordogne se recogieron 500 encuestas. Las fichas se distribuyeron en cuatro grupos: 120 niños de zona urbana, 120 niñas de zona urbana, 130 niñas de zona rural y 130 niñas de zona rural. En el Département de la Marne 173 escolares realizaron la prueba: dos grupos de 75 niños y 98 niñas, de los cuales provenían 133 de ciudad y 40 del campo. En el Département de l'Eure la muestra contó 85 informantes, formados por 73 niños y 12 niñas, todos (salvo 5) de la ciudad. Por último, en el Département de la Vendée se pasaron 146 cuestionarios realizados a 90 niños y 56 niñas, 130 correspondientes a la ciudad y 16 al campo.

Centro	Estudio actual	Départments de la Dordogne, Marne, Eure, Vendée	Département de la Dordogne
01. Le corps humain	tête, main, bras, pied, bouche, oreille, jambe, doigt, nez, œil, yeux, cheveu, cou, ongle, dent, ventre, dos, épaule, genou, langue	oeil (yeux), oreille, nez, bras, jambes, pied, main, bouche, tête, cœur, doigts, cou, foie, estomac, dents, poumons, langue, cheveux, genoux, ventre	

02. Les vêtements	pantalon, chaussure, jupe, chemise, t(ee)-shirt, robe, chaussette, veste, manteau, jean, gant, chapeau, pull(-over), écharpe, cravate, culotte, blouse, foulard, collant, botte	chemise, chaussettes, culotte, manteau, robe, veste, pull-over, jupe, blouse, gilet, tablier, pantalon, cravate, cache-nez, combinaison, béret, chapeau, caleçon, bas, corsage	chemise, veste, tablier, manteau, gilet, culotte, robe, chaussette, jupe, combinaison, pantalon, cravate, bas, gants, béret, imperméable, blouse, chapeau, corsage, pull-over
06. Moyens de transport	voiture, vélo, avion, train, bus, métro, bateau, tram(way), cheval, moto, TGV, taxi, à pied, RER, TER, camion, tricycle, navette, parachute, minibús		camion, train, avion, bateau, moto, auto, charrette, car, bicyclette, micheline, autobus, tramway, vélo, vélomoteur, camionnette, brouelle, tracteur, diligenc, remorque, paquebot

Tabla 4.25. Comparación de los veinte primeros vocablos en inglés de las investigación de Gougenheim *et al.* (1964) y del estudio actual

Esta investigación solo permite hacer varias comparaciones ya que no recoge el listado de palabras ordenadas según el índice de disponibilidad léxica. No obstante, se pueden observar algunos datos llamativos. Por una parte, se reitera el carácter compacto del primer centro puesto que el vocabulario presente en varios departamentos en Francia y en la muestra actual coincide en un 85 %. A continuación, se comprueba que en el centro referido a la vestimenta, coinciden los tres listados en un 55 % pero el resto de palabras escritas por franceses y españoles solo sumaría un 5 %. Sin embargo, las dos últimas columnas añaden un 30 % al porcentaje de coincidencia donde se incluyen términos como *combinaison* y *coursage*, que quizás hoy en día no serían tan disponibles.

CI	Tres investigaciones		Estudio actual y varios departamentos		Estudio actual y Dordogne		Varios departamentos y Dordogne	
	Vocablo	%	Vocablo	%	Vocablo	%	Vocablo	%
1	-	-	17	85 %	-	-	-	-
2	11	55 %	0	-	1	5 %	6	30 %
6	-	-	-	-	8	40 %	-	-

Tabla 4.26. Datos comparativos de la investigación de Gougenheim *et al.* (1964) y del estudio actual

En líneas generales, se ha comprobado que los listados se asemejan más o menos según se trate de la lengua materna o la lengua extranjera de los informantes. Este análisis inicial podría dar lugar a un estudio más pormenorizado de la organización del lexicón mental como se realiza en el Capítulo 6. Además, es posible conocer más allá de las palabras localizadas, como puede ser que ciertas profesiones estén más relacionadas con un sexo u otro, razón por la cual se dedica el último

apartado de este capítulo a presentar los listados de palabras según el sociolecto. Además, en lengua extranjera es posible conocer la presencia del léxico en los manuales o la facilidad de adquirir ciertos términos.

4.6. La situación ortográfica

En el presente estudio, la recogida de errores ortotipográficos de cada informante se realizó mediante una matriz diseñada *ad hoc*. El formato seguido para la confección del listado fue el mismo que para la lematización de las palabras, por lo que se creó una hoja de cálculo por centro de interés donde a cada informante le correspondía una fila que incluía diferentes tipos de anotaciones. Los errores detectados han sido registrados en su forma original, es decir, sin aplicar las normas de edición. Por tanto, se recogen palabras en género femenino, *arquitecta tecnica*, y número plural, *arboles*²¹⁷.

A continuación, se presenta la distribución de disortografías por idioma y por centros de interés. El objetivo es recoger qué tipo de faltas de ortografía son las más acuciadas para reflexionar sobre las posibles causas y buscar estrategias de enseñanza y aprendizaje adecuadas que ayuden a soslayar ese obstáculo. Además, se realiza una comparación intermuestral con resultados de otros estudios que recogen los errores ortográficos de sus listados desde el punto de vista descriptivo (Paredes García, 1999; Saura Rami, 2008; Blanco Domínguez, 2011) y sociolingüístico (Ávila Muñoz, 2007).

4.6.1. Resultados del estudio actual

La presentación de las palabras concretas con distortografías se ha incluido en forma de apéndice separados según el idioma: español, inglés y francés. En español se presenta según el modelo recogido en el denominado «Apéndice ortográfico» del estudio realizado en Málaga por Ávila Muñoz y Villena Ponsoda (2010: 317-319):

- Acentuación

²¹⁷La conservación de las formas originales permite la realización de estudios ulteriores en español y en lengua extranjera centrados en la didáctica de la lengua. En este sentido, cabe mencionar el análisis de los errores léxicos en español como lengua extranjera de los estudiantes finlandeses realizado por Carcedo González (1999c).

- Grafía c, z + e, i
- j-g
- s-x-c-k-qu
- h
- b-v
- ll-y-ñ
- Otras

En cambio, el formato de presentación del anexo en las lenguas extranjeras difiere del español debido a la dificultad de lematizar algunos vocablos que se alejan de la forma correcta. Ejemplo de ello, es el informante 25, cuando trata de escribir *caracol* en inglés, *snail*, pero lo hace en francés, *scargot*. Por ello, los errores ortográficos y calcos de otros idiomas van precedidos de la forma correcta en la que se han lematizado los vocablos.

Del total de vocablos con faltas de ortografía, es decir 328, más de la mitad corresponden a errores de acentuación, concretamente 217 vocablos. El resto de disortografías son menores, con lo que desde el punto de vista pedagógico parece necesario insistir en la aplicación de las reglas de acentuación. Cabe destacar que en algunas programaciones de centros educativos las faltas de ortografía penalizan la calificación del alumnado. Por tanto, el profesor ha de ser capaz de corregir adecuadamente las producciones escritas del alumnado y convertirse en un referente para el mismo.

CI	Acentuación	c, z + e, i	j-g	s-x-c-k-qu	h	b-v	ll-y-ñ	Otras	Total
C1	35	0	1	0	0	3	2	5	46
C2	13	0	1	1	0	3	1	9	28
C3	28	1	2	0	2	3	0	15	51
C4	22	0	2	1	0	1	0	8	34
C5	29	1	1	1	0	0	0	5	37
C6	19	1	0	2	0	0	0	8	30
C7	24	0	0	0	0	0	1	15	40
C8	29	0	0	2	0	1	0	5	37
C9	18	0	0	0	0	2	0	5	25
Total	217	3	7	7	2	13	4	75	328

Tabla 4.27. Clasificación de errores ortográficos en lengua española

La Tabla 4.27 corrobora la hipótesis de estudios anteriores y muestra una gran escala entre las faltas de ortografía debidas a la acentuación gráfica, un 66,16 %, respecto al resto de errores, 33,84 %. El apartado «Otras» recoge todas aquellas disortografías que no han sido incluidas en el resto de secciones, entre las que destacan las faltas relacionadas con palabras extranjeras como *leggings* y *paddle*.

La distinción entre pares de letras que designan un mismo fonema, tales como *b* y *v* que corresponden al fonema /b/ y las letras *j* y *g* en los casos que se refieren al fonema /x/, presentan igualmente ciertas dificultades en los hablantes. Un 3,95 % de los errores se derivan del uso de *b* y *v* que corresponden a trece vocablos (*bello*, *púvis*, *vazo*, *baquero*, *verenjena*, *vikini*, *aluvia*, *cerbeza*, *vicarbonato*, *bate*, *avogado*, *gravador*, *gravadora*) y un 2,13 % al uso de la *j* y *g* relativo a siete vocablos (*vajina*, *trage*, *berengena*, *potage*, *conserge*, *juete*, *agetreo*).

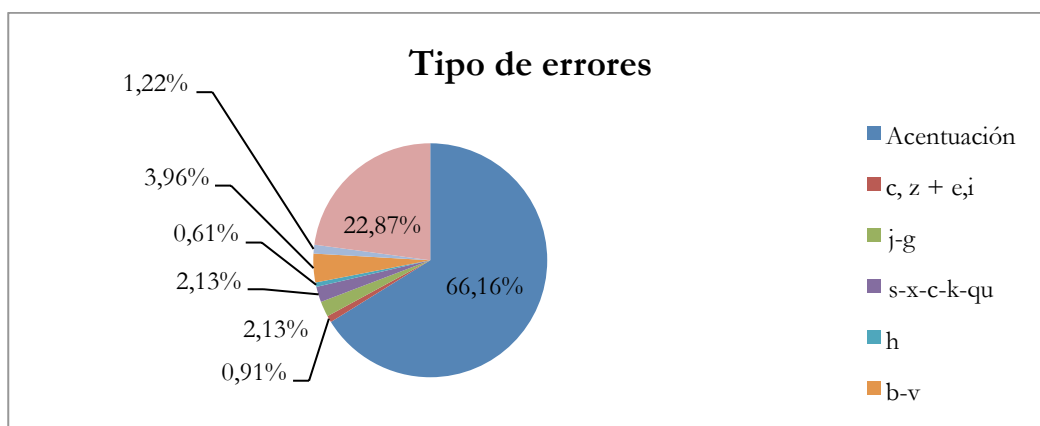


Gráfico 4.12. Distribución porcentual del tipo de errores en lengua española

En el Gráfico 4.13 se muestra la distribución porcentual de los errores ortotipográficos y de las palabras en cada centro de interés. En general la diferencia entre las palabras y errores no supera el 1 % por ámbito, es decir, que coincide el porcentaje de palabras en un centro de interés con respecto al total. Por ejemplo, el primer centro, *Partes del cuerpo humano*, cuenta con un 13,69 % de palabras y un 13,39 % de los errores totales. No obstante, el séptimo centro, *Juegos y distracciones*, presenta una diferencia de algo más de cuatro puntos debido quizás a que en ese apartado se han incluido una gran cantidad de nombres comunes y propios de juegos, que a menudo están presentes en el lenguaje oral pero no son tan frecuentes en un texto escrito. Además, ese centro de interés contiene un 13,39 % de errores al igual que el primer centro y se encuentra por debajo del tercer centro, *Comidas y bebidas*. Sin

embargo, estos dos últimos centros son considerablemente más productivos, 13,69 % y 14,32 % frente a 9,26 % del séptimo centro, de ahí que el porcentaje de faltas de ortografía se corresponda con el número de palabras producidas.

En el lado opuesto se sitúan el segundo centro, *La ropa*, y el noveno, *Ordenadores e internet*, en cuyos casos el porcentaje de errores es considerablemente inferior al de palabras, 11,45 % del total de palabras en comparación con 9,47 % del total de errores en el primer caso y 10,03 % frente al 7,62 % en el segundo. Tanto en el segundo como en el noveno centro, existe una cantidad considerable de extranjerismos que han sido acuñados y forman parte de nuestro repertorio lingüístico, tales como, *anorak*, *pashmina*, *hardware* y *software*, entre otras.

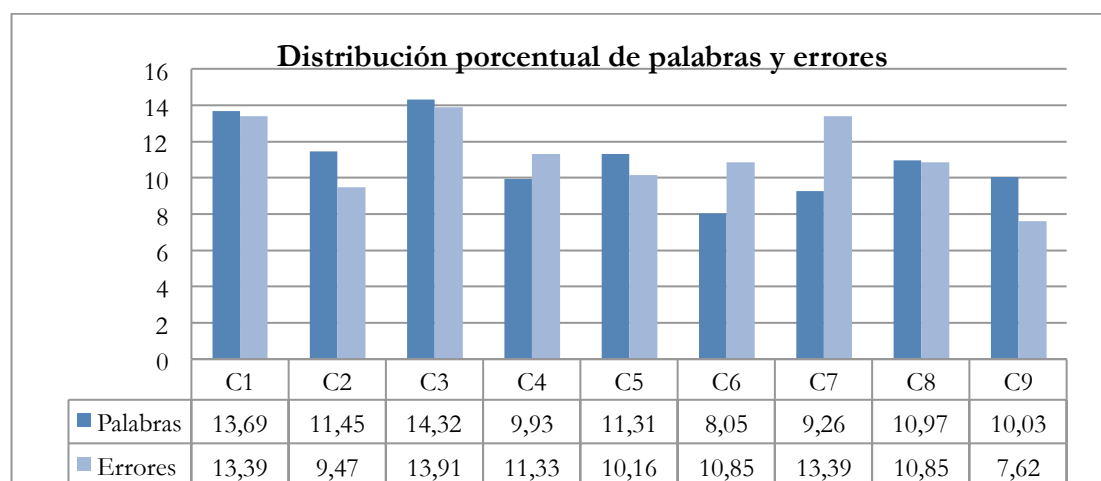


Gráfico 4.13. Distribución porcentual de errores y palabras por centros de interés en español

La Tabla 4.28 recoge el número de palabras total aportadas por centro de interés, el número de palabras con errores y la relación porcentual entre ambas. El porcentaje de lexías incorrectas oscila entre el 1 % del centro 9, *Ordenadores e internet*, hasta el 1,91 % del centro 7, *Juegos y distracciones*. Aunque *a priori* el número parezca muy reducido, el profesor es el modelo de escritura del alumnado y cometería un error por cada 100 palabras, por ejemplo, el equivalente a una falta en cada seis o siete líneas de este trabajo.

De igual modo se ha presentado la relación de vocablos con errores respecto a su número total. En este caso las diferencias entre los distintos centros de interés se acentúa. El primer centro de interés, *Partes del cuerpo humano*, muestra el mayor índice de vocablos con errores con un 18,85 % ya que de 244 lexías diferentes 46 han sido escritas de forma incorrecta. Sin embargo, el centro con más faltas de ortografía es el

tercero, *Comidas y bebidas*, con un total de 51 pero debido a su gran productividad solo presenta un 9,62 % de sus vocablos incorrectos. Ello se debe a que el centro 1 muestra un vocabulario más homogéneo que el resto²¹⁸ con un índice de cohesión de 0,108 mientras que el índice del centro 3 desciende hasta 0,050 y, por tanto, en el primer caso la proporción es mayor que en el segundo.

CI	Total palabras	Palabras incorrectas	%	Total vocablos	Vocablos incorrectos	%
C1	4493	58	1,29	244	46	18,85
C2	3759	41	1,09	238	28	11,76
C3	4565	62	1,36	530	51	9,62
C4	3718	43	1,16	392	34	8,67
C5	3711	44	1,19	483	37	7,66
C6	2641	47	1,78	190	30	15,79
C7	3040	58	1,91	469	40	8,53
C8	3600	47	1,31	478	37	7,74
C9	3290	33	1,00	555	25	4,50
Total	32817	433	1,32	3579	328	9,16

Tabla 4.28. Palabras y vocablos incorrectos con respecto al total en español

En lengua extranjera la distribución porcentual de palabras por centros de interés es similar al español, a excepción del centro 9, *Ordenadores e internet*, donde aumenta el porcentaje de palabras y vocablos aportados en lengua extranjera con respecto al español. Los errores ortográficos se concentran en mayor medida en el centro de interés 8, *Profesiones y oficios*, con un 19,29 % del total de disortografías aunque paradójicamente es el penúltimo centro en porcentaje de palabras con un 9,58 %. Existen palabras como *engineer* cuyas disortografías han adoptado hasta diez formas diferentes y otras muy frecuentes como, *professor*, que presentan un gran número de errores repetidos en tres variantes: *profesor*, *proffessor* y *proffesor*.

En porcentaje de faltas de ortografía le sigue el centro 3, *Comidas y bebidas*, con un 17,95 %, lo cual está justificado por su peso en el porcentaje de palabras aportadas, el mayor con un 14,63 %. El centro con mejores datos ortotipográficos es el cuarto, *La escuela: muebles y materiales*, debido a que suele ser el tema más recurrido en clase al aprender un idioma extranjero. Este hecho puede considerarse positivo ya que el futuro profesorado que pretenda impartir clases bilingües deberá conocer y escribir adecuadamente todo el vocabulario relacionado con su entorno escolar.

²¹⁸ La información detallada sobre el índice de cohesión de cada centro se encuentra en el apartado 4.3.

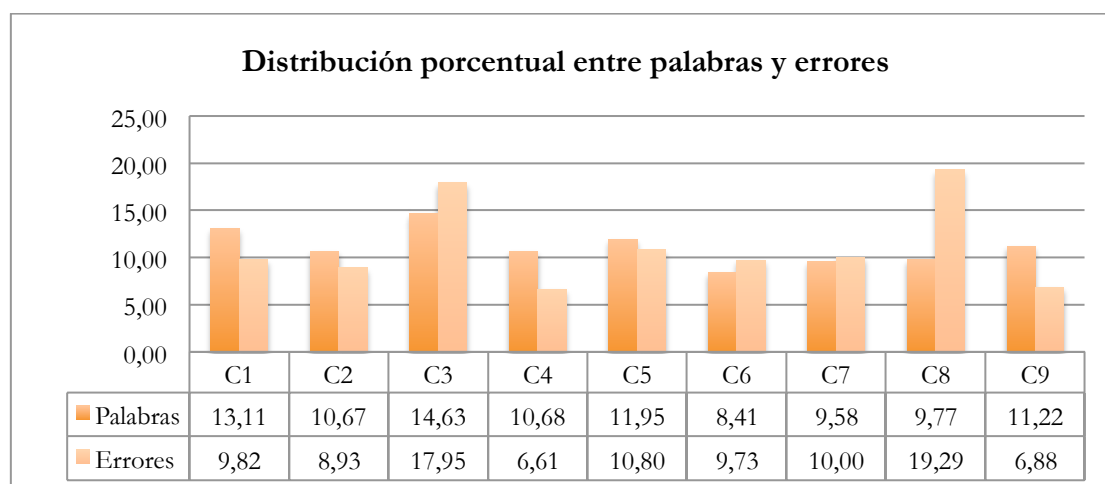


Gráfico 4.14. Distribución porcentual de errores y palabras por centros de interés en lengua extranjera

A continuación, se presenta la relación entre las palabras y vocablos con errores ortográficos con respecto al total aportado por los hablantes en inglés y en francés. Mientras que de cada 100 palabras en español los hablantes escriben una media de 1,32 palabras incorrectas, en inglés es de 4,46 y en francés asciende a 7,67. Al igual que en español, la mayoría de los errores en francés se deben a la acentuación incorrecta de las palabras. Además, en francés el acento ortográfico adopta tres formas diferentes: acento circunflejo, acento grave y acento agudo. Esta dificultad añadida provoca en cierta medida que el número de palabras incorrectas sea mayor que en inglés.

El centro con mayor porcentaje de errores es el octavo, *Profesiones y oficios*, donde 9,04 palabras de cada 100 están escritas de forma incorrecta. El centro que le sigue, *Comidas y bebidas*, presenta una mejora considerable ya que desciende a un 5,37 % de palabras incorrectas. El centro con menor porcentaje de errores es el noveno, *Ordenadores e internet*, debido en gran parte al uso del vocabulario relacionado con las nuevas tecnologías en forma de préstamo en español, lo que muestra una gran familiarización con la escritura de las palabras.

CI	Total palabras	Palabras incorrectas	%	Total vocablos	Vocablos incorrectos	%
C1	2571	88	3,42	115	35	30,43
C2	2076	88	4,24	151	35	23,18
C3	2894	167	5,77	287	69	24,04
C4	2120	50	2,36	244	35	14,34
C5	2344	85	3,63	343	53	15,45
C6	1675	90	5,37	165	31	18,79
C7	1909	90	4,71	310	48	15,48
C8	1928	174	9,02	280	77	27,5

C9	2251	50	2,22	460	41	8,91
Total	19768	882	4,46	2355	424	18

Tabla 4.29. Palabras y vocablos incorrectos con respecto al total en inglés

Como ocurre en lengua española, el primer centro, *Partes del cuerpo humano*, es el más compacto y, por tanto, los vocablos son poco numerosos en comparación con otros centros. Por tanto, el porcentaje de vocablos con errores es superior al del resto de centros de interés ya que han sido mencionados de forma más frecuente. En el caso opuesto se sitúa el centro de interés relacionado con las nuevas tecnologías que presenta un total de 460 vocablos y tan solo 41 con errores, por lo que el porcentaje descende al mínimo 8,91 %.

Respecto a la lengua francesa, el porcentaje por errores en cada centro supera los registrados en español y en inglés. Por un lado, el centro que ocupa el primer rango coincide con el inglés, siendo el octavo, *Profesiones y oficios*, con un 14,05 %. No obstante, en este caso el centro con menor número de palabras incorrectas con respecto a su total es el segundo, *La ropa*, con 3,29 palabras de cada 100 palabras aportadas.

Por otro lado, el centro con menor proporción y número de vocablos incorrectos es el cuarto referido a la escuela, con ocho vocablos y un 7,08 %. En cambio, el sexto centro presenta mayor porcentaje, un 20,75 %, ya que los informantes tan solo aportaron 53 vocablos diferentes y se presentan con 11 disortografías. En cuanto al noveno centro, *Ordenadores e internet*, la proporción de vocablos con faltas de ortografía se sitúa en torno a la media con un 14,17 %, puesto que el inglés es el idioma de referencia en el campo de las nuevas tecnologías en detrimento del francés y otras lenguas que han incorporado las voces inglesas a ese dominio léxico.

CI	Total palabras	Palabras incorrectas	%	Total vocablos	Vocablos incorrectos	%
C1	427	22	5,15	78	16	20,51
C2	365	12	3,29	80	14	17,50
C3	452	34	7,52	166	22	13,25
C4	323	24	7,43	113	8	7,08
C5	389	36	9,25	135	23	17,04
C6	249	19	7,63	53	11	20,75
C7	282	22	7,80	125	14	11,20
C8	299	42	14,05	119	21	17,65
C9	316	27	8,54	127	18	14,17
Total	3102	238	7,67	996	147	14,76

Tabla 4.30. Palabras y vocablos incorrectos con respecto al total en francés

Resulta igualmente interesante comprobar qué porcentaje de informantes ha realizado la prueba de disponibilidad léxica sin cometer ningún error ortográfico. En la Tabla 4.31 se observa que en español se contabilizan, en este sentido, 36 informantes mientras que en lengua extranjera tan solo se documentan 6 informantes. A pesar de ello, puede resultar preocupante que tan solo el 21,05 % del futuro profesorado sea capaz de realizar una prueba de léxico con una ortografía impecable.

Español			Lengua extranjera		
Informantes	Faltas	Porcentaje	Informantes	Faltas	Porcentaje
36	0	21,05	6	0	3,51
31	1	18,13	14	1	8,19
19	2	11,11	9	2	5,26
19	3	11,11	17	3	9,94
11	4	6,43	23	4	13,45
17	5	9,94	11	5	6,43
1	6	0,58	11	6	6,43
5	7	2,92	22	7	12,87
4	8	2,34	12	8	7,02
3	9	1,75	15	9	8,77
4	10	2,34	8	10	4,68
4	11	2,34	2	11	1,17
1	12	0,58	3	12	1,75
2	13	1,17	5	13	2,92
6	14	3,51	2	14	1,17
2	16	1,17	3	15	1,75
3	17	1,75	2	17	1,17
1	20	0,58	1	18	0,58
1	24	0,58	1	19	0,58
1	33	0,58	1	20	0,58
			1	28	0,58
			1	29	0,58

Tabla 4.31. Distribución de informantes según el número de faltas de ortografía en español y en lengua extranjera

En español y en lengua extranjera el número de informantes que ha cometido más de 10 faltas de ortografía es muy similar: 21 y 22 informantes respectivamente. La diferencia en ambas lenguas radica en los informantes con cinco o menos faltas de ortografía: en español asciende a 133 informantes y en lengua extranjera tan solo llega a 80. Los estudiantes con mayor número de faltas de ortografía en español han

escrito en lengua española 33 palabras de forma incorrecta y en lengua extranjera 29 ya que como norma general, han sido menos productivos en lengua extranjera.

4.6.2. Resultados comparativos intermuestrales

Los estudios de disponibilidad léxica que analizan la ortografía de las palabras aportadas seleccionan a informantes preuniversitarios o universitarios (Paredes García, 1999; Ávila Muñoz, 2007; Saura Ramí, 2008; Blanco Domínguez, 2011; García Casero, 2013). De esta forma, es posible valorar la progresión de forma cuantitativa de varios estudios con estudiantes preuniversitarios y los postgraduados del estudio actual. No obstante, el hecho de que los resultados no se presenten de igual forma en los diferentes análisis impedirá el cotejo simultáneo de todas las investigaciones.

Paredes García (1999) aplica la prueba de disponibilidad seleccionando diecisiete centros, de los cuales dieciséis corresponden a los tradicionales y uno es denominado *Los colores*. La muestra escogida es de más de 400 alumnos de 3.º y 4.º de Educación Secundaria Obligatoria y de 1.º y 2.º de Bachillerato de Alcalá de Henares.

Con el propósito de realizar un análisis comparativo, se ha calculado la media de palabras por informante y centros de interés. Se observa una diferencia considerable entre la situación ortográfica del alumnado preuniversitario y de postgrado ya que, en el primer caso 11,34 de cada 100 palabras contienen alguna errata y en el segundo tan solo 1,32 palabras. Además, la producción léxica entre los estudiantes postgraduados es superior a la de los preuniversitarios: cada informante de postgrado escribe de media 21,32 palabras por centro de interés mientras que el estudiante preuniversitario proporciona 18,94 unidades léxicas. Por tanto, el futuro profesorado no solo aporta mayor cantidad de palabras sino que también lo hace de forma más ajustada a la norma ortográfica.

En el Apartado 4.4, se muestra que el caudal léxico en lengua materna siempre es superior al de una lengua extranjera. Así, el promedio de unidades léxicas por informante es mayor en la prueba en español de los estudiantes de postgrado (21,32) seguida de la realizada con los estudiantes preuniversitarios (18,94) y de la prueba en lengua extranjera (14,86). En cambio, la corrección ortográfica atiende a otros

factores. En el caso del estudio de Paredes García (1999), de cada 100 palabras 11,34 % contiene una errata ortográfica mientras que en la presente investigación con estudiantes de postgrado, en español el porcentaje desciende 1,32 % y en lengua extranjera aumenta al 4,90 %. Los datos arrojados muestran que el idioma seleccionado influye el nivel de instrucción del alumnado pero, sobre todo, incide el nivel de estudios alcanzados por los informantes.

	Preuniversitarios madrileños	Postgraduados (español)	Postgraduados (lengua extranjera)
Porcentaje palabras ortográficamente incorrectas	11,34 %	1,32 %	4,90 %
Media de errores ortográficos por sujeto y centro	2,26	0,28	0,73
Media palabras por sujeto y centro	18,94	21,32	14,86

Tabla 4.32. Tabla comparativa de los errores ortográficos del estudio de Paredes García (1999) y la muestra del presente estudio en español y en lengua extranjera

En el artículo de Ávila Muñoz (2007) se analiza la corrección ortográfica de 50 estudiantes de segundo curso de la Licenciatura de Traducción e Interpretación de la Universidad de Málaga sobre los 16 centros de interés tradicionales. En el estudio no se recogen las palabras concretas escritas de forma incorrecta ya que el análisis se centra en determinar los condicionamientos sociales que influyen en la ortografía de la prueba.

En la Tabla 4.33, se observa que los resultados de los estudiantes postgraduados presentan mayor porcentaje en los dos primeros apartados, referidos a ningún error y menos de diez errores lo cual supondría un 85,38 % de la muestra en lengua española y un 81,87 % en lengua extranjera. En cambio, en el estudio de Ávila Muñoz (2007) tan solo un 34 % de los informantes se sitúa entre los dos primeros parámetros. Igualmente es significativa la distribución de informantes en el último nivel referido a más de 20 errores ortográficos siendo del 28 % en el estudio de los estudiantes universitarios y de 1,17 % en el caso de los estudiantes postgraduados.

	Universitarios malagueños		Postgraduados (español)		Postgraduados (lengua extranjera)	
Nº de faltas	Informantes	Porcentaje	Informantes	Porcentaje	Informantes	Porcentaje
0 faltas	3	6 %	36	21,05 %	6	3,51 %
Menos de 10 faltas	14	28 %	110	64,33 %	134	78,36 %
Entre 10 y 20 faltas	19	38%	23	13,45 %	29	16,96 %
Más de 20 faltas	14	28 %	2	1,17 %	2	1,17 %

50	171	171
----	-----	-----

Tabla 4.33. Tabla comparativa de los errores ortográficos del estudio de Ávila Muñoz (2007) y la muestra del presente estudio en español y en lengua extranjera

Saura Rami (2008) selecciona 45 informantes de las encuestas de disponibilidad realizadas todas las provincias aragonesas a estudiantes de 2.º de Bachillerato o COU. Al igual que en el artículo de Paredes García (1999) los centros de interés seleccionados corresponden a los 16 de los tradicionales más uno adicional denominado *Los colores*.

De la Tabla 4.34 se desprende que los estudiantes gallegos son más productivos que los posgraduados malagueños pero también que su ortografía es considerablemente más descuidada. Las media de palabras aportadas por los estudiantes de postgrado en español es de 21,32, dato similar al promedio de 23,59 del estudio de Saura Rami. El porcentaje de faltas de ortografía es superior no solo al de la prueba en lengua española sino también a los resultados en lengua extranjera siendo de 7,72 % en el estudio actual frente al 4,90 % de los aragoneses. No obstante, en comparación con otros léxicos de estudiantes preuniversitarios, tales como la investigación de Paredes García (1999) con un 11,34 %, los estudiantes aragoneses presentan una buena corrección léxica con un porcentaje de errores ortográficos menor, 7,72 %.

	Preuniversitarios aragoneses	Postgraduados (español)	Postgraduados (lengua extranjera)
Porcentaje palabras incorrectas	7,72 %	1,32 %	4,90 %
Media faltas por sujeto y centro	1,83	0,28	0,73
Media palabras por sujeto y centro	23,59	21,32	14,86

Tabla 4.34. Tabla comparativa de los errores ortográficos del estudio de Saura Rami (2008) y la muestra del presente estudio en español y en lengua extranjera

El análisis llevado a cabo por Blanco Domínguez (2011) se nutre de las encuestas realizadas a informantes preuniversitarios gallegos sobre 18 centros de interés: 16 correspondientes a los clásicos más *Los colores* y *El mar*. Los datos presentados en el estudio recogen los errores ortográficos de las palabras que presentan un índice de frecuencia acumulado del 80 %. De las 243 241 palabras diferentes, la autora detecta 11 152 errores, lo que equivale a un 4,59 % del total. Porcentaje superior al estudio actual en lengua española (1,32 %) y en lengua extranjera (4,90 %).

Para concluir, cabe mencionar que de los análisis en lengua española y en lengua extranjera se deduce que el número de errores de un centro de interés dado no está directamente relacionado con el número de palabras aportadas. En algunos supuestos, como el caso del noveno centro en lengua extranjera, *Ordenadores e internet*, se ha comprobado que el porcentaje de errores es inferior al porcentaje de palabras producidas, 6,88 % frente al 11,22 %. En cambio, en otros centros surge el fenómeno contrario, es decir, que los informantes han escrito un menor número de palabras pero con una mayor cantidad de faltas ortográficas como es el caso del octavo centro en lengua extranjera, *Profesiones y oficios*, con 19,29 % de faltas y 9,77 % de palabras.

En cuanto a los análisis intermuestrales, se constata que el porcentaje de errores de los estudiantes postgraduados en lengua española es inferior al de los estudiantes universitarios y preuniversitarios. Asimismo, salvo en el análisis contrastivo con el estudio de Saura Rami (2011) el porcentaje de errores en lengua extranjera es menor que los resultados en lengua española de los estudiantes preuniversitarios y universitarios por lo que el factor con mayor incidencia es el nivel de estudios alcanzados por los informantes.

4.7. Incidencia de la variable ‘sexo’ en el léxico disponible

Al igual que Gómez Molina (2006: 64), en este apartado se abordan los resultados del sociolecto femenino en contraposición con el masculino desde una perspectiva cualitativa y cuantitativa. En el estudio han participado un total de 54 hombres y 117 mujeres. Por ello, teniendo en cuenta que no existe paridad en la distribución del número de informantes por sexo, se comparan datos que no se ven influidos por el número de hablantes, como el promedio de palabras por sujeto y centro de interés.

En la segunda columna de la Tabla 4.35 se ha recogido el número de vocablos que solo están presentes en uno de los listados que contienen los 20 primeros vocablos más disponibles, de tal forma que si coinciden todos a excepción de uno, como en el primer centro, se registra con el número 1. En la tercera columna, se han registrado los vocablos aportados por los hombres y a la derecha los vocablos

aportados por las mujeres seguidos del total de lexías diferentes en ese centro. La cuarta columna se divide también en tres para dar cuenta de los vocablos aportados por los hombres, por las mujeres y la suma de ambos. Por último, la quinta, recoge el porcentaje de vocablos diferentes entre un sexo y otro.

El centro más homogéneo es el denominado *Partes del cuerpo humano* y, por ende, solo se detecta un vocablo en las primeras 20 posiciones, que no coincide en los dos listados. Asimismo, se observa que el total de lexías diferentes aportadas por hombres y mujeres es muy similar, 189 y 209, y el número de vocablos aportados por los dos subgrupos es de tan solo 90 palabras, 35 los hombres y 55 las mujeres dando lugar a un porcentaje de 36,89 % de vocablos diferentes. Un comportamiento similar se observa en el sexto centro, *Medios de transporte*, donde 18 de los 20 vocablos más disponibles coinciden y los vocablos diferentes recogidos por hombres y mujeres son 93.

El centro más heterogéneo es el noveno, *Ordenadores e internet*, que cuenta con un total de 555 vocablos. Por ello, el léxico aportado por los informantes masculinos y femeninos es muy diverso: los hombres han escrito 152 vocablos que no aparecen en los listados de las mujeres y ellas 203, dando lugar a un total de 355 lexías exclusivas de un sociolecto u otro y el porcentaje más alto: 63,96 %.

CI	Diferencia entre los 20 primeros	Total vocablos			Vocablos diferentes			Vocablos comunes
		Hombres	Mujeres	Ambos	Hombres	Mujeres	Ambos	
C1	1	189	209	244	35	55	90	36,89 %
C2	4	150	205	238	33	88	121	50,84 %
C3	5	349	431	530	99	181	280	52,83 %
C4	2	222	325	392	67	170	237	60,46 %
C5	4	281	404	483	79	202	281	58,18 %
C6	2	140	147	190	43	50	93	48,95 %
C7	4	268	387	469	82	201	283	60,34 %
C8	4	319	388	478	90	159	249	52,09 %
C9	4	352	403	555	152	203	355	63,96 %
Media	3,33	252,22	322,11	397,67	75,56	145,44	221	53,84 %

Tabla 4.35. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en español

En inglés, la distribución de los vocablos por centros de interés es similar al español. El centro más compacto es el primero y, por ello, es el que cuenta con menos unidades léxicas diferentes entre hombres y mujeres, tan solo 45. Asimismo, el centro más heterogéneo es el noveno, con un total de 305 palabras exclusivas de un centro u otro. En términos porcentuales, el vocabulario diferente de ambos centros

sitúa en un extremo el primer centro, *Human body parts*, con un 39,13 % de palabras diferentes y el último, *Computer and internet*, con un 66,30 %.

CI	Diferencia entre los 20 primeros	Total vocablos			Vocablos diferentes			Vocablos comunes
		Hombres	Mujeres	Ambos	Hombres	Mujeres	Ambos	
C1	1	81	105	115	11	34	45	39,13 %
C2	4	98	130	151	21	54	75	49,67 %
C3	4	194	227	287	60	93	153	53,31 %
C4	4	144	203	244	41	100	141	57,79 %
C5	6	198	277	343	66	145	211	61,52 %
C6	1	104	123	165	42	61	103	62,42 %
C7	6	186	251	310	59	124	183	59,03 %
C8	5	166	229	280	51	114	165	58,93 %
C9	5	269	346	460	114	191	305	66,30 %
Media	4	160	210,11	261,67	51,67	101,78	153,45	56,46 %

Tabla 4.36. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en inglés

En francés, los datos difieren con respecto al español e inglés. La muestra está compuesta por un número inferior, 21 informantes, de ahí que en la primera columna los resultados sean superiores a los otros dos idiomas. En este caso, los informantes han tenido menos oportunidades para homogeneizar el vocabulario y, por tanto, la disponibilidad léxica de los vocablos solo se ha calculado a partir de una veintena de respuestas.

Asimismo, algunas de las palabras aportadas por un solo grupo podrían encuadrarse en el otro grupo si contasen con un número superior de sujetos. El centro con menor porcentaje de diferencia entre los vocablos aportados por hombres y mujeres es el séptimo, *Jeux et loisirs*, con un 18,40 %, seguido del quinto, *La ville*, con un 20,74 %. En los centros más productivos como es el caso del tercer y cuarto centro, la diferencia es considerablemente mayor, 62,65 % y 68,14 %, en consonancia con el elevado número de vocablos 166 y 113.

CI	Diferencia entre los 20 primeros	Total vocablos			Vocablos diferentes			Vocablos coincidentes
		Hombres	Mujeres	Ambos	Hombres	Mujeres	Ambos	
C1	2	49	71	78	7	29	36	46,15 %
C2	8	52	65	80	15	28	43	53,75 %
C3	7	94	134	166	32	72	104	62,65 %
C4	7	49	100	113	13	64	77	68,14 %
C5	10	81	28	135	28	54	82	20,74 %
C6	5	40	15	53	15	13	28	28,30 %
C7	10	64	23	125	23	62	85	18,40 %
C8	8	70	29	119	30	49	79	25,21 %
C9	11	75	33	127	33	52	85	25,98 %
Media	7,56	63,78	53,33	110,67	55,56	47	102,56	38,81 %

Tabla 4.37. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en francés

En cuanto al promedio de vocablos aportados por hombres y mujeres, en todos los casos el sociolecto femenino presenta un mayor caudal léxico debido a que el número de informantes hombres es inferior al de mujeres, por lo que esos datos no se han considerados representativos de la muestra. En cambio, es posible realizar la comparación de la media de palabras aportadas según sexo. En este caso, la media masculina en español solo supera a la femenina en el noveno centro, *Ordenadores e internet*. El número de informantes hombres de la especialidad de Tecnología, Informática y Procesos Industriales es superior al de las mujeres, por lo que muestran mayor dominio del caudal léxico en el centro referido a los ordenadores.

CI	Media de vocablos diferentes en español			Media de palabras en inglés			Media de vocablos en francés		
	Hombres	Mujeres	Ambos	Hombres	Mujeres	Ambos	Hombres	Mujeres	Ambos
C1	24,93	26,9	26,28	16,13	17,51	17,14	17,38	22,77	20,33
C2	19,91	22,96	21,98	12,89	14,21	13,84	15,75	18,77	17,38
C3	25,02	27,47	26,7	18,3	19,68	19,29	20,25	22,69	21,52
C4	20,56	22,3	21,74	13,98	14,15	14,13	13,25	17,08	15,38
C5	20,89	22,08	21,7	15,09	15,76	15,63	17,5	20	18,52
C6	15,22	15,55	15,44	11,35	11,05	11,17	12,25	11,92	11,86
C7	17,19	18,05	17,78	12	12,99	12,73	12,5	14,46	13,43
C8	20,26	21,42	21,05	12,76	12,84	12,85	13,5	15,15	14,24
C9	20,06	18,86	19,24	15,54	14,68	15,01	14,25	16,23	15,05
Media	20,45	21,73	191,91	128,04	132,87	131,79	136,63	159,08	147,71

Tabla 4.38. Media de vocablos en los diferentes idiomas según sexo

Aunque las divergencias en lengua extranjera no son estadísticamente significativas, se presentan a modo de ejemplo la media de palabras por informantes en lengua extranjera para comprobar si se repite el mismo modelo que en español. En este caso, también los hombres superan la media de las mujeres en dos centros de interés: en el de *Ordenadores e internet*, como ocurre en los resultados en español, y en *Medios de transporte*.

Al realizar un análisis de correlaciones bivariadas con el número de lexías aportadas por informante en cada centro y la variable sexo en el programa SPSS, se constata que las diferencias solo son significativas en tres centros de interés. La correlación es significativa al nivel 0,01 (bilateral) en el caso de *La ropa* y al nivel 0,05 en los centros *Comidas y bebidas*, *La escuela: muebles y materiales*. En el diagrama de cajas se muestra la diferencia entre las medias de ambos subgrupos en el centro 2, *La ropa*, siendo superior la aportada por las mujeres (M). Se observan dos casos extremos de las encuestadas identificadas como 21 y 34 con un número de palabras muy por

encima del promedio. La muestra de los informantes masculinos es más homogénea y, por tanto, la distancia entre los valores mínimos y máximos es más acortada.

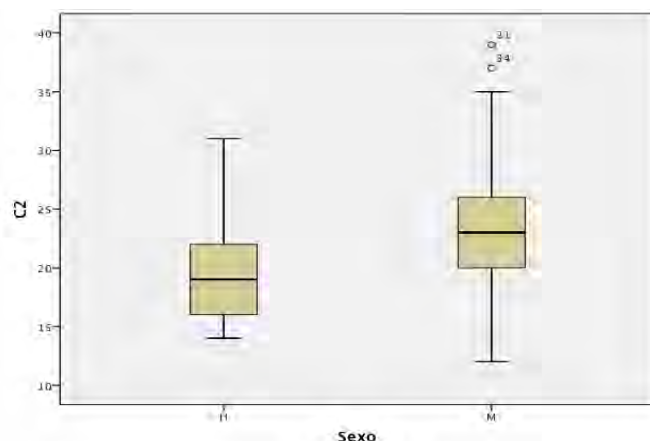


Gráfico 4.15. Diagrama de cajas que recoge la distribución por sexos del número de palabras del centro denominado *La ropa*

Dada la relevancia de la variable sexo en tres centros de interés, se presentan los veinte vocablos más frecuentes aportados por los hombres y mujeres de la población de estudio para comprobar el grado de coincidencia. Asimismo, se analizan las unidades léxicas que han sido aportadas solo por el subgrupo femenino o masculino. Para el resto de centros de interés, la información se encuentra disponible en el Anexo IX.

Gómez Molina (2006: 64) concluye que la diferencia entre vocablos y palabras aportados por el sociolecto femenino y masculino viene marcada biológica y culturalmente. En este sentido Bellón Fernández (2011: 119) destaca la idiosincrasia de los centros de interés y de su relación con el sexo de los informantes:

Por tanto, tendremos que concluir que los resultados obtenidos en nuestro estudio del léxico disponible de la provincia de Córdoba no son fortuitos y que efectivamente nos hallamos ante un hecho general: algunos centros de interés son más favorables para unos hablantes que para otros, dependiendo de su sexo. Aún más, esas áreas temáticas - manejadas en los estudios de disponibilidad léxica- más propicias para las mujeres suelen ser ‘La ropa’, ‘Los muebles de la casa’, ‘Alimentos y bebidas’, ‘Objetos colocados en la mesa para la comida’, ‘La cocina y sus utensilios’ y ‘Los colores’; mientras que los campos léxicos habitualmente más accesibles a los hombres son ‘Calefacción, iluminación y medios de airear un recinto’, ‘El campo’, ‘Medios de transporte’ y ‘Trabajos del campo y del jardín’.

4.7.1. Centro 2: *La ropa*

Las 20 palabras más disponibles proporcionadas por hombres y mujeres coinciden en más de un 75 %, concretamente en 16 vocablos. Las lexías diferentes en el grupo masculino son *sombrero*, *cinturón*, *gorra* y *corbata* que se asocian a prendas de vestir tradicionalmente más usuales entre los hombres que entre las mujeres. Del mismo modo, el grupo femenino aporta los vocablos *vestido*, *guante*, *rebeca* y *bota*. Los sustantivos referidos a la ropa interior, *braga* y *calzoncillo*, corresponden a un par de palabras que los informantes escriben en uno u otro orden hasta en 51 ocasiones, de ahí que aunque se trate de prendas de mujer y hombre aparezcan como vocablos disponibles en los dos listados. Del mismo modo, se observa que algunos vocablos compartidos por ambos grupos cuentan con un porcentaje de aparición mayor o menor según el sexo. Por ejemplo, *falda* ha sido escrita por el 61,11 % de los hombres y por el 85,47 % de las mujeres.

Hombres			Mujeres		
Vocablos	ID	% Aparición	Vocablos	ID	% Aparición
pantalón	0,81101	92,593	pantalón	0,79874	96,581
camiseta	0,67209	85,185	falda	0,65699	85,47
camisa	0,626	75,926	camisa	0,61125	81,197
calcetín	0,54719	81,481	jersey	0,59933	77,778
chaqueta	0,53423	72,222	camiseta	0,5914	76,923
zapato	0,50085	74,074	zapato	0,5223	82,051
jersey	0,4829	62,963	vestido	0,51243	73,504
falda	0,41389	61,111	chaqueta	0,47651	70,94
calzoncillo	0,36523	57,407	bufanda	0,44305	69,231
braga	0,3177	51,852	calcetín	0,43887	76,923
abrigo	0,27886	42,593	abrigo	0,39764	64,103
bufanda	0,26667	46,296	media	0,31582	56,41
sombrero	0,25353	37,037	guante	0,28677	49,573
cinturón	0,24557	40,741	rebeca	0,26422	41,026
gorra	0,24167	38,889	braga	0,26202	49,573
media	0,23859	40,741	gorro	0,24553	43,59
sujetador	0,22879	37,037	sujetador	0,24098	43,59
(pantalón) vaquero	0,2284	33,333	(pantalón) vaquero	0,23829	37,607
gorro	0,21637	33,333	bota	0,23604	47,009
corbata	0,20402	37,037	calzoncillo	0,21359	43,59

Tabla 4.39. Veinte primeros vocablos del centro 2, *La ropa*, según la variable sexo

En cuanto a los vocablos aportados exclusivamente por hombres o mujeres, se destacan los siguientes aspectos:

- Los hombres incluyen prendas de vestir femeninas tales como *body*, *burkini*, *sostén*, *trikini* y *Wonderbra*.

- Las mujeres aportan vocablos más concretos sobre el tipo de ropa: *bota alta, bota de media caña, bota de montaña, falda de tubo, falda de vuelo, falda pantalón, traje de baño, traje de noche, traje de vestir, vestido de flamenca, vestido de noche, vestido de novia, vestido de tirantes, pañuelo palestino, pashmina*.
- Los vocablos referidos a complementos se encuentran más presentes en el léxico del sociolecto femenino: *accesorio, alianza, antifaz, bisutería, brazalete, broche, felpa, gargantilla, joya, muñequera*. En el sociolecto masculino: *anteojo, complemento, peluca, reloj de pulsera*.

Hombre	Mujer
anteojo, arreglarse, babucha, body, burkini, camisola, capucha, chamarra, complemento, conjunto, estilo casual, formal, informal, jeans, lencería, manga a la sisa, mitón, moda, pantalón de pana, pantufla, peluca, probador, reloj de pulsera, ropa blanca, ropa de color, sobrecamisa, sombrero chambergo, sostén, tienda, triquini, turbante, Wonderbra, zanco	(pantalón) pirata, abalorio, accesorio, alianza, antifaz, bailarina, bastón, bata de cola, bisutería, blazer, bolsa, bolsillo, Borsalino, bota alta, bota de media caña, bota de montaña, braga (del cuello), brazalete, broche, calza, calzado, camiseta de tirantes, casaca, chamarreta, chaqué, colgante, colorido, combinación, corsé, cremallera, delantal, deportivo, diminuto, disfraz, ejecutivo, equipación, estola, faja, fajín, falda de tubo, falda de vuelo, falda pantalón, felpa, forro, gargantilla, gomilla, hebilla, jersey de cuello alto, joya, lana, manga, mantilla, mantón, media tupida, mono, muñequera, niqui, orejera, palabra de honor, pana, pantalón de pinzas, pañuelo palestino, pashmina, percha, picardía, pichi, ropa, ropa deportiva, sombrero de copa, talla grande, talla pequeña, tejano, tirante, toalla, tocado, torera, traje de baño, traje de noche, traje de vestir, túnica, uniforme, vestido camisero, vestido de flamenca, vestido de noche, vestido de novia, vestido de tirantes, zapatilla de casa, zueco

Tabla 4.40. Vocablos aportados solo por hombres y mujeres en el centro 2, *La ropa*

4.7.2. Centro 3: *Comidas y bebidas*

Este centro de interés presenta un repertorio lingüístico más variado que el anterior ya que el índice de cohesión es de 0,050 mientras que en *La ropa* es de 0,092. El grado de coincidencia entre los veinte primeros vocablos de los dos listados es el que presenta la mayor diferencia de todos los centros con 5 palabras, lo que corresponde a un 25 %. Sin embargo, esas lexías no podrían considerarse características de un sociolecto u otro ya que aparecen en posiciones más avanzadas y con un alto índice de disponibilidad. Cabe destacar la alta disponibilidad del vocablo *café* en las mujeres (0,237) con respecto a los hombres (0,127) quizás por referirse, más que a la bebida, a la costumbre de reunirse con otras personas en la sobremesa.

Hombres			Mujeres		
Vocablos	ID	% Aparición	Vocablos	ID	% Aparición
cerveza	0,33273	48,148	agua	0,497	66,667
agua	0,32011	46,296	tomate	0,375	64,103
pescado	0,31738	51,852	zumos	0,366	53,846

carne	0,30879	46,296	vino	0,332	48,718
tomate	0,28071	53,704	carne	0,332	46,154
vino	0,27126	42,593	pan	0,321	49,573
patata	0,26556	46,296	patata	0,313	52,137
arroz	0,25628	44,444	pescado	0,313	48,718
Coca-Cola	0,24428	33,333	cerveza	0,311	45,299
lechuga	0,23401	40,741	leche	0,309	47,009
ensalada	0,22796	37,037	lechuga	0,294	52,137
zumos	0,22585	33,333	café	0,237	35,043
leche	0,22201	29,63	arroz	0,235	39,316
pan	0,22095	37,037	verdura	0,231	37,607
fruta	0,18611	33,333	naranja	0,217	39,316
pollo	0,18089	29,63	fruta	0,208	29,915
manzana	0,1678	29,63	refresco	0,204	29,915
tortilla	0,15807	22,222	pasta	0,201	31,624
pizza	0,15728	27,778	Coca-Cola	0,196	25,641
huevo	0,15596	22,222	manzana	0,188	34,188

Tabla 4.41. Veinte primeros vocablos del centro 3, *Comidas y bebidas*, según la variable sexo

Este centro cuenta con 530 lexías diferentes y se sitúa en el segundo rango tras el centro 9, *Ordenadores e internet*, con 555 vocablos. Dada la heterogeneidad del léxico, la diferencia de las palabras aportadas solo por informantes masculinos y femeninos se acentúa, 280 vocablos. Los hombres aportaron 99 lexías que no aparecen en los listados de las mujeres y ellas a su vez aportaron 181 unidades léxicas diferentes. Tras analizar los vocablos en sendos sociolectos, se pueden destacar las siguientes singularidades:

- Los hombres tienden a escribir más palabras relacionadas con bebidas alcohólicas: *absenta, aguardiente, albariño, Bacardi, Cacique, Chardonnay, cubalibre, Dyc, ginebra, JB, Martini, pacharán, ponche, Ribera del Duero, sake, sauvignon, Sex on the Beach, vino blanco y vino rosado*. En cambio, las mujeres solo recogen en exclusividad los vocablos genéricos: *alcohol, brandy y sidra*.
- Las mujeres han aportado más vocabulario relacionado con postres y alimentos dulces: *arroz con leche, caramelo, chuchería, chupachús, cuajada, flan, flan de huevo, gofre, golosina, gominola, graniñada, horchata, leche merengada, muffin, napolitana, natilla, pionono, tocino de cielo, tofu y vainilla*. Los hombres solamente incluyen: *churro, repostería, sirope, tarta de chocolate y tarta de fruta*.

Hombre			Mujer
absenta,	aguardiente,	ajeno,	aceituna rellena, agua con gas, agua mineral, alcohol, alioli, alita

albariño, anguila, aperitivo, azúcar moreno, Bacardi, bebida, bistec, boca de mar, bombón, boquerón en vinagre, Cacique, canapé, cangrejo, caña, cardo, carne de cerdo, caviar, cazuela de fideos, chacina, Chardonnay, chipirón, choco, chopo, churro, ciervo, cigala, conserva animal, conserva vegetal, coquina, criadilla, cubalibre, Danone, Doritos, Dyc, espeto, fiambre, filete tártaro, fusilli, gallina, gallo, gamba rebozada, ginebra, guarnición, gula, hojaldre, jamón cocido, JB, kiko, La Casera, láudano, lechal, manita de cerdo, Martini, mate, mero, mijo, mojito, mozzarella, nabo, naranjada, nugget, pacharán, palmera, papaya, pero, pimiento morrón, pinchito, piquito, ponche, porra, potaje de acelgas, potaje de lentejas, quesadilla, queso manchego, repostería, revuelto de espárragos, Ribera del Duero, rioja, rollito de primavera, rollo, sacarina, sake, sauvignon, secreto ibérico, seso asado, Sex on the Beach, sirope, sofrito, tarta de chocolate, tarta de fruta, tintorería, vino blanco, vino rosado, zumo de arándano, zumo de melocotón, zumo de piña	de pollo, almuerzo, anacardo, Aquarius, arándano, arroz con leche, asado, avellana, baguette, banana, barra de pan, bebida alcohólica, bebida isotónica, beicon, berberecho, berenjena con miel, bollo, brandy, brécol, brote de soja, buey, burrito, caballa, cacahuete, cacao, calamarito, caldo Avecrem, caldo gallego, calzone, camarón, canela, caramelo, carne a la plancha, carne mechada, carpaccio, cazuela, cebolleta, cena, chicle, chuchería, chupachús, chupito, cochinillo, cocinar, coco, comida basura, concha fina, condimento, conejo a la cazadora, costilla, crema pastelera, cuajada, cuchara, cuchillo, cuscús, desayuno, dorada, empanada, ensalada de arroz, entrante, escarola, Fanta Limón, Fanta Naranja, fideuá, flan, flan de huevo, fréjol, fruto rojo, gazpachuelo, gelatina, gofre, golosina, gominola, granizada, hambre, hielo, hierbabuena, hígado, horchata, huevo cocido, huevo de codorniz, ingrediente, jabalí, jengibre, jibia, kebab, leche merengada, lomo en manteca, macedonia, manteca, mantecado, mantel, maracuyá, McDonald's, melocotón en almíbar, merluza, migas, mitad, mollete, mora, muffin, muslo, napolitana, natilla, nube, ñoqui, ostra, palmito, palomita, pan cateto, pan de ajo, pan de molde, panceta, papilla, pasa, patata a lo pobre, pechuga, pepinillo, Pepsi, picadillo, pimentón, piñón, pionono, pipa de calabaza, pipa de girasol, pisto, pitufo, pollo al curry, pollo asado, producto biológico, Puleva, puré, queso blanco, queso brie, queso de untar, rábano, rabo de toro, rebanada, rebujito, receta, Red Bull, restaurante, revuelto, revuelto de verdura, riñón al jerez, ruca, salami, salmonete, salsa boloñesa, salsa de yogur, salsa rosa, salteado, shawarma, sidra, sobrasada, soda, solomillo a la pimienta, sopa de marisco, sopa de picadillo, Sprite, tapa, té verde, tenedor, tocino, tocino de cielo, tofu, tomate cherry, tomillo, tortellini, tortilla francesa, tostada, vaca, vainilla, valeriana, wrap, zurrapa
---	---

Tabla 4.42. Vocablos aportados solo por hombres y mujeres en el centro 3, *Comidas y bebidas*

4.7.3. Centro 4: *La escuela: muebles y materiales*

La media de palabras aportadas por los hombres en este centro es de 20,56 lexías mientras que el promedio de las mujeres es de 22,30. El número de vocablos aportados por ambos grupos se sitúa en torno a la media (397,67) con 392 vocablos. Solo se destacan dos lexías diferentes en los listados que incluyen los 20 primeros vocablos: *puerta* y *alumno* en el caso de los hombres; *rotulador* y *carpeta* en el listado de las mujeres. Al igual que en los centros anteriores, ese vocabulario está presente en los dos listados pero con un índice de disponibilidad inferior.

Hombres			Mujeres		
Vocablos	ID	% Aparición	Vocablos	ID	% Aparición
pizarra	0,78054	90,741	pizarra	0,73421	87,179
mesa	0,70113	83,333	silla	0,6945	82,906
silla	0,65629	81,481	lápiz	0,63974	91,453
boli(grafo)	0,58154	81,481	mesa	0,63962	71,795
lápiz	0,56808	77,778	boli(grafo)	0,57956	84,615
tiza	0,53632	72,222	libro	0,471	75,214
libro	0,42431	72,222	goma (de borrar)	0,44064	68,376

pupitre	0,39129	46,296	pupitre	0,41867	48,718
goma (de borrar)	0,34779	53,704	tiza	0,41593	55,556
regla	0,27677	44,444	ordenador	0,41377	66,667
profesor	0,26359	37,037	libreta	0,32291	51,282
cuaderno	0,23179	37,037	estuche	0,28855	52,137
estuche	0,22473	38,889	borrador	0,27566	39,316
borrador	0,2223	31,481	rotulador	0,27243	46,154
libreta	0,21876	37,037	regla	0,26967	45,299
ordenador	0,21639	33,333	cuaderno	0,23811	40,171
puerta	0,20883	29,63	mochila	0,22625	41,88
alumno	0,20247	29,63	profesor	0,20467	27,35
mochila	0,19944	40,741	proyector	0,19224	29,915
proyector	0,19619	31,481	carpeta	0,17137	33,333

Tabla 4.43. Veinte primeros vocablos del centro 4, *La escuela: muebles y materiales*, según la variable sexo

Las columnas de la Tabla 4.44 recogen un total de 237 palabras diferentes: 67 corresponden al grupo masculino y 170 al grupo femenino. Entre los vocablos aportados se observa la segregación de sexos en la enseñanza en centros religiosos ya que los hombres han escrito *cura*, los cuales suelen enseñar en colegios para chicos, mientras que las mujeres han escrito *monja*, quienes están más presentes en colegios para chicas.

En cuanto al vocabulario específico, el subgrupo masculino ha escrito más nombres de asignaturas: *ciencias naturales, educación física, educación plástica y visual, filosofía, francés, geografía, geología, griego, inglés, latín, lengua castellana y literatura, lenguaje y tecnología* mientras que las mujeres no han mencionado ninguna asignatura, solo el sustantivo genérico: *materia*. Por su parte, en el listado femenino se localizan los nombres de las diferentes partes de un centro educativo y los tipos de centro: *aula de convivencia, campus, centro bilingüe, centro TIC, colegio, edificio, escuela, escuela concertada, escuela privada, escuela pública, guardería, hall, institución, instituto, jefatura de estudios, polideportivo, sala de informática, sala de música, seminario y universidad*; en cambio, los hombres solo han escrito *cantina*.

Hombre	Mujer
alfombra, alumnado, ascensor, auricular, baldosa, Biblia, bigotera, biología, boca de incendio equipada, bombilla, butaca, cajonera, campana, cantina, celofán, chándal, ciencias naturales, cuadrícula, cura, curso académico, diario, dintel, educación física, educación plástica y visual, esquema, estante, filosofía, forro, francés, fútbol, geografía, geología, globo	(rotulador) fluorescente, abrigo, administrativo, afilador, agua, aire acondicionado, amigo, AMPA, aprender, aprendizaje, aprobado, aprobar, archivador, aro, atlas, atril, aula de convivencia, babero, barro, bloc, bloc de dibujo, bloc de notas, bola del mundo, bolso, cable, caja, caligrafía, cama, campus, canasta, cancha de baloncesto, capilla, cartilla, cartuchera, CD(-ROM), celo, centro bilingüe, centro TIC, chincheta, colegio, columpio, comportamiento, concurso, conocimiento, consejo escolar, copistería, cortina, cuadro, cuerda, desayuno, dirección, diversión, dossier, edificio, enemigo, enseñar, escalímetro, escáner, escuela, escuela concertada, escuela privada, escuela pública, espejo, estudiante, estudio, ética, excursión, fax, Fixo, fotocopia, fuente, funda, grande, grapa, grupo, guardería, guitarra, hall, iluminación, imagen, impresora, institución, instituto, jardín, jefatura de estudios, juego,

terráqueo, griego, imbornal, inglés, latín, lengua castellana y literatura, lenguaje, manualidad, mapamundi, marcador, matrícula, panel, pantalla digital, papel cebolla, papel vegetal, parte, pendrive, pizarra blanca, potro, producto químico, profesorado, rampa para minusválidos, religión, reloj, ropero, tarea, tecnología, teléfono, tenis, test, trabajo, transportador, transporte escolar, tubo fluorescente, urinario	juguete, lámina, lápiz de madera, lavabo, lectura, librería, libro de lectura, libro de texto, libro del profesor, líquido corrector, madrugar, manual, materia, mesa de dibujo, método, mina, monja, muñeco, mural, netbook, niño, ordenador personal, orientador, página, pandereta, papel de seda, papel milimetrado, papel pinocho, paragüero, parte de faltas, pegatina, película, personal de limpieza, pica, Pilot, pincel, pizarra tradicional, plumier, polideportivo, porta-ángulos, portafolio, portería, portero, probeta, programación, pupilo, quitagrapas, radio, raqueta, ratón, redacción, repaso, reproductor, reproductor de DVD, reproductor de vídeo, respaldo, resumen, reunión, revista, sala de informática, sala de música, sanción, secretario, semicírculo graduado, seminario, Séneca, separador, servicio, suspender, suspenso, tablón de corcho, taller, tambor, teclado, televisor, temario, transportador de ángulos, trimestre, tutor, unidad didáctica, uniforme, universidad, váter, vídeo
---	---

Tabla 4.44. Vocablos aportados solo por hombres y mujeres en el centro 4, *La escuela: muebles y materiales*

Debido a la disparidad entre el número de informantes de sexo masculino y femenino, resulta evidente que las mujeres hayan aportado más vocablos que los hombres. No obstante, el promedio de palabras por centro de interés sitúa al sociolecto femenino casi siempre por encima del masculino. Aunque se realizase un estudio sobre qué factores inciden en el sexo y viceversa, es decir, si resulta que las mujeres tienden a escoger carreras universitarias del área lingüística mientras que los hombres del área científica y que esa variable a su vez índice en el léxico disponible, el sexo no deja constancia de diferencias estadísticamente significativas en la mayoría de los centros²¹⁹.

4.8. Conclusiones parciales

La lengua materna de los informantes resulta ser la más productiva en cuanto a número de palabras y vocablos. En español, los informantes han escrito un promedio de 191,91 palabras, en inglés 131,79 y en francés 147,71. En cuanto a los vocablos, en español se registra una media de 20,93, en inglés de 13,77 y en francés de 5,82. Estos datos tan diferenciados se debe a que los vocablos se ven influenciados por el número de informantes: 171 en español, 150 en inglés y 21 en francés.

Al comparar los resultados con otros estudios realizados en Málaga, concretamente sobre una muestra de estudiantes preuniversitarios (Ávila Muñoz,

²¹⁹ En el capítulo siguiente no se incluye la variable «sexo» porque en los análisis estadísticos no mostró una influencia significativa sobre las variables analizadas.

2006) y de población adulta (Ávila Muñoz y Villena Ponsoda, 2010) se destaca la mayor productividad de los estudiantes de postgrado. La media de palabras por centro de interés calculada sobre los ocho centros comunes, es de 19,69 en los estudiantes preuniversitarios, 20,50 en la población adulta estratificada y 21,59 en los estudiantes de postgrado. Ello indica que *a priori*, el nivel de estudios parece tener cierta influencia en el léxico disponible que se analizará con más detenimiento en el capítulo siguiente.

En los estudios de español como lengua extranjera se ha comparado la media de esta investigación en lengua extranjera ya que como es obvio los resultados en español eran considerablemente superiores. Los estudiantes finlandeses (Carcedo González, 2000c) han aportado una media de 13,68 palabras por centro de interés, los extranjeros que realizan una estancia en Salamanca (Samper Hernández, 2002) superan esa media con un promedio 15,63 palabras y los españoles del estudio actual, 17,24. Por tanto, el nivel alcanzado en lengua extranjera por los estudiantes españoles es considerablemente más alto que en el resto de casos.

En inglés, la media de palabras es inferior a los estudios con hablantes nativos. Para los centros presentes en la investigación de Ferreira (2006) se ha recalculado la media de este estudio, 23,19 palabras, mientras que los nativos encuestados por el investigador chileno han escrito 27,28. Del mismo modo, los escolares de la investigación de Dimitrijević (1969) aportan una media de 20,04 palabras en inglés y, en cambio, la media de los estudiantes de postgrado en ese idioma es de 17,69 palabras. No obstante, cabe resaltar que el promedio de respuestas de los estudiantes de postgrado españoles es superior al de los estudiantes universitarios chilenos, con un promedio de 23,19 frente a 20,33 palabras por informante y centro de interés. Aunque los informantes en el estudio de Chile no han acabado la carrera universitaria, todos cursan estudios relacionados con el aprendizaje del inglés como lengua extranjera por lo que quizás los resultados deberían ser más altos.

En los casos anteriores, se ha analizado la amplitud léxica. En cambio, la propuesta de este estudio radica en comparar la media de palabras según el índice corregido de amplitud léxica (ICoAL) ya que tiene en cuenta no solo el número de palabras aportadas por un informante, sino también la relación con el centro de

interés. De esta forma, se podrá evaluar de forma más rigurosa la capacidad léxica de los informantes. No obstante, habría que establecer un consenso en las palabras consideradas nucleares (N1), asociaciones colectivas (N5), derivaciones (N25) y asociaciones individuales (N0). Asimismo, esta nueva forma de contabilizar el léxico favorece la inclusión de todo tipo de palabras en los listados, sobre todo de las palabras fruto de asociaciones individuales ya que aparecerán en los últimos puestos del listado del léxico disponible, no serán contabilizadas en el ICoAL pero servirán de gran ayuda para conocer la organización de las palabras en la mente de los informantes.

En cuanto al índice de cohesión, se observa una tendencia común en las tres lenguas que se refleja en los rangos de los centros según su índice de cohesión y densidad léxica. El centro más cohesionado es el primero, *Partes del cuerpo humano*, en los tres idiomas mientras que los menos cohesionados son *Juegos y distracciones* y *Ordenadores e internet*. Al comparar la clasificación realizada con otros estudios, tales como, en Ávila Muñoz y Villena Ponsoda (2010) se reitera que el centro más homogéneo es el primero. Por tanto, en el capítulo 6, se analizará más en profundidad la estructura interna de los centros de interés presuponiendo que el tipo de estímulo marca las respuestas de los informantes.

En los análisis de los veinte vocablos más disponibles, se ha corroborado la hipótesis que presupone una mayor similitud en los listados de estudios de una lengua extranjera con respecto a estudios en lengua materna. El uso y el contexto de aprendizaje influyen directamente en el léxico disponible. De esta forma, entre los vocablos más disponibles en lengua materna los informantes han mencionado cuatro relacionadas con la ropa interior mientras que en lengua extranjera solo han aportado una o dos. Ese tipo de prendas se usan diariamente pero no aparece con tanta frecuencia en los manuales de inglés, francés o español como lengua extranjera. Asimismo, el espacio temporal de recogida de datos muestra la diacronía de la lengua condicionando el vocabulario más disponible, razón por la cual en el estudio pionero de Gougenheim *et al.* (1964) aparecen vocablos como *combinaison* que probablemente hoy en día no serían tan disponibles porque ha cambiado la forma de vestir.

Con respecto al análisis ortográfico, los errores se han clasificado en ocho tipos: 1) acentuación; 2) grafía c, z + e, i; 3) j-g; 4) s-x-c-k-qu; 5) h; 6) b-v; 7) ll-y-ñ; 8) otras. El 66,16 % de las disortografías corresponden a una inadecuada acentuación de las palabras, ya sea porque se ha acentuado cuando no debiere o viceversa. Además del grupo que engloba faltas de ortografías varias, con un porcentaje de 3,96 % se sitúa el uso incorrecto de los grafemas *b* y *v* ya que la poligrafía del fonema /b/ provoca confusiones. En este sentido, le siguen los grupos de faltas de ortografía 3 y 4. Estos resultados arrojan datos de gran interés en el ámbito educativo ya que muestra que habría que incidir sobre todo en las reglas de acentuación. Además, cabría plantearse alguna medida correctora en la universidad o en niveles inferiores ya que el futuro profesorado no debería tener dificultades en escribir palabras tan disponibles como *bikini*²²⁰.

Sin embargo, en líneas generales se comprueba que los estudiantes de postgrado han cometido menos errores que los estudiantes preuniversitarios, dando así cuenta de su paso por la universidad. Mientras que los preuniversitarios madrileños (Paredes García, 1999) cometen una media de 2,26 faltas de ortografía por informante y centro de interés y los preuniversitarios aragoneses 1,83 (Saura Rami, 2008), los estudiantes de postgrado tan solo 0,28 en español y 0,73 en lengua extranjera. En cuanto al porcentaje de errores de los universitarios malagueños (Ávila Muñoz, 2007) se observa que es considerablemente superior al de los postgraduados. Por ejemplo, en el primer caso solo existe un 6 % de informantes que no presentan disortografía alguna frente al 21,05 % de los estudiantes de postgrado en español. En cambio, en lengua extranjera ese porcentaje desciende al 3,51 %.

Por último, se ha analizado el repertorio lingüístico del sociolecto femenino y masculino por tratarse de la variable más universal que está presente en la mayoría de los estudios, desde el pionero realizado por Gougenheim *et al.* (1964) hasta nuestros días. En general, las diferencias entre los listados de hombres y mujeres son poco relevantes (Gougenheim *et al.*, 1956; Dimitrijević, 1969; Echeverría *et al.*, 1987; Alba Ovalle, 1995a; Mateo García, 1998; Valencia y Echeverría, 1991; Carcedo González, 2001) y, en algunas ocasiones, el centro de interés condiciona el mayor caudal léxico

²²⁰ En el Anexo VII se ha incluido el resumen de todos los vocablos con faltas de ortografía.

de un sexo u otro. Aunque en la mayoría de los centros de interés, el sociolecto femenino parece ser más productivo no se detectan diferencias estadísticamente significativas salvo en los centros siguientes: *La ropa y Comidas y bebidas*, *La escuela: muebles y materiales*. Lagüéns Gracia (2008:139) y Trigo Ibáñez y González Martínez (2011: 39) recomiendan un análisis más exhaustivo de la incidencia de la variable ‘sexo’ con respecto a otras variables y, que en este caso, forma parte del cometido del próximo capítulo.

CAPÍTULO 5. ANÁLISIS SOCIOLINGÜÍSTICO

5.1. Introducción

5.2. Hipótesis

5.3. Análisis descriptivo univariante y presentación de las variables

5.3.1. Variables lingüísticas

5.3.1.1. Índice corregido de descentralización léxica

5.3.1.2. Índice corregido de identificación de tecnicismos (ICIT)

5.3.2. Variables sociodemográficas y reticulares

5.3.2.1. Edad

5.3.2.2. Estudios

5.3.2.3. Ocupación de los padres e ingresos

5.3.3. Variables relacionadas con el dominio de una lengua

5.3.3.1. Especialidad del máster

5.3.3.2. Conocimiento de otras lenguas extranjeras

5.3.3.3. Uso de la lengua extranjera en diferentes contextos

5.3.3.4. Exposición a los medios de comunicación y libros al año en español

5.3.3.5. Motivación del futuro profesorado

5.3.4. Resumen de las variables

5.4. Análisis bivariante

5.4.1. Relaciones entre variables lingüísticas

5.4.1.1. Variables lingüísticas en lengua extranjera

5.4.1.2. Variables lingüísticas en lengua española

5.4.1.3. Caudal léxico en lengua extranjera y en lengua española

5.4.1.4. Léxico especializado en lengua extranjera y en lengua española

5.4.2. Efecto de las variables socioculturales

5.4.2.1. Estudios del padre y léxico especializado en lengua extranjera

5.4.2.2. Estudios del padre y caudal léxico en lengua española

5.4.3. Efecto de las variables relacionadas con el dominio de una lengua

5.4.3.1. Especialidad del máster

5.4.3.2. Conocimiento de idiomas

5.4.3.3. Uso de la lengua extranjera en diferentes contextos

5.4.3.4. Hábitos de lectura y exposición a los medios de comunicación en español

5.4.3.5. Motivación para impartir clases en enseñanza bilingüe

5.5. Análisis multivariante

5.5.1. Análisis de componentes principales (ACP). Reducción de dimensiones

5.5.1.1. Clase social

5.5.1.2. Motivación para impartir clases bilingües

5.5.1.3. Uso de la lengua en diferentes ámbitos

5.5.2. Análisis de regresión lineal

5.5.2.1. Análisis de regresión lineal del caudal léxico en lengua extranjera

5.5.2.2. Análisis de regresión lineal del caudal léxico en español

5.5.2.3. Análisis de regresión lineal del léxico especializado en lengua extranjera

5.5.2.4. Análisis de regresión lineal del léxico especializado en lengua española

5.5.3. Análisis causales

5.5.3.1. Modelo causal de los efectos del hablante sobre el caudal léxico en lengua extranjera

5.5.3.2. Modelo causal de los efectos del hablante sobre el caudal léxico en lengua española

5.5.3.3. Modelo causal de los efectos del hablante sobre el léxico especializado en lengua extranjera

5.5.3.4. Modelo causal de los efectos del hablante sobre el léxico especializado en lengua española

5.6. Conclusiones parciales

5.1. Introducción

Según Fernández Fernández, Cordero Sánchez y Córdoba Largo (2002: 15-16) la estadística ha evolucionado desde las primeras civilizaciones, donde se recogían datos sobre todo con fines recaudatorios, hasta nuestros días por medio de las ciencias básicas. Los orígenes de esta disciplina están relacionados con dos necesidades: registrar información y conocer la probabilidad de éxito en los juegos de azar. Ambos menesteres serán el punto de partida de las dos grandes ramas de la estadística: la descriptiva y la inferencial. *Grosso modo*, la estadística descriptiva se encarga de ordenar, presentar y reducir los datos recopilados de una población o muestra. En cambio, la estadística inferencial, también conocida como estadística inductiva o deductiva, se basa en la teoría de probabilidades para hacer predicciones sobre los datos.

A lo largo de este capítulo se llevarán a cabo indagaciones sobre todo de tipo descriptivo completadas con estudios inferenciales como el Análisis de Anova imprescindibles para conocer el comportamiento de distintas variables. De esta forma, se pretende ordenar la información relativa a las variables extralingüísticas y lingüísticas generadas a partir de las respuestas proporcionadas en las encuestas de disponibilidad léxica, el cuestionario sociológico y la prueba de léxico especializado. Cabe destacar que la base de datos está formada por más de 500 variables que permiten un profundo conocimiento de la muestra de estudio y, por ende, de sus características socioculturales, reticulares y académicas, entre otras. Por ello, se han seleccionado aquellos factores que contribuyen a refutar o corroborar las hipótesis de la investigación.

Antes de iniciar los análisis, se detallan las hipótesis que marcan la línea de trabajo y que intentan dar respuesta a los interrogantes del estudio. En un primer estadio, se presentan las variables a través de análisis univariantes haciendo referencia al proceso de recogida y codificación de los datos. A continuación, los análisis bivariantes muestran la relación entre dos características de la muestra y ayudan a definir los análisis posteriores. Por último, los análisis multivariantes presentan una

visión multidimensional para explicar la contribución de factores sociales, lingüísticos y relacionados con el aprendizaje de una lengua extranjera sobre el índice corregido de descentralización léxica (ICoDD) y el índice corregido de identificación de tecnicismos (ICIT).

5.2. Hipótesis

Se parte de la hipótesis que el comportamiento sociolingüístico y los antecedentes relacionados con el aprendizaje de una lengua inciden en el vocabulario de los informantes. El tipo de estudios previos, la clase social, los hábitos de lectura y el uso de la lengua extranjera ayudan a predecir los resultados del léxico disponible y la capacidad de identificación de tecnicismos. Por una parte, el caudal léxico es medido a través de las pruebas de disponibilidad léxica en español y en lengua extranjera (inglés o francés), las cuales aportan información sobre el número de palabras que ha actualizado un hablante en un tiempo determinado y sobre la relación que establecen con el centro de interés. Por otra parte, el léxico especializado se obtiene mediante un ejercicio que consiste en identificar y relacionar a qué campo léxico corresponden palabras especializadas o técnicas. El objetivo es analizar de forma global el vocabulario activo y el vocabulario pasivo de un informante sobre nueve centros de interés.

La probabilidad de que la influencia de ciertos factores releven mejores resultados en el léxico viene supeditada por la lengua de estudio. Así, se presupone que no tendrá igual relevancia la realización de una estancia en el extranjero en lengua extranjera que en la lengua materna. En español, Ávila Muñoz y Villena Ponsoda (2010) destacaron el papel de las redes sociales y la posición social sobre la capacidad léxica. En cambio, en lengua extranjera los estudios existentes no realizan una profundización tan precisa del conjunto multifactorial que determina el léxico disponible y la capacidad de identificación del léxico especializado.

En lengua materna, las hipótesis apuntan a que el vocabulario activo y, sobre todo, el vocabulario pasivo están condicionados por el nivel de instrucción. En este estudio, todos los informantes son al menos diplomados, licenciados o ingenieros ya que es uno de los requisitos previos de acceso al máster. A pesar de esta aparente

homogeneización en el grado de formación de los informantes, esta variable adquiere matices a través de otros factores como la carrera universitaria, la especialidad del máster y otros estudios complementarios.

Entre las variables con mayor incidencia en el léxico, cabe destacar el conocimiento de idiomas extranjeros. Como se abordó en los fundamentos teóricos, el dominio de más de una lengua aporta agilidad mental y permite establecer conexiones entre una mayor cantidad de unidades léxicas que presumiblemente influye en el mayor caudal léxico²²¹. Por consiguiente, el léxico disponible en español y en lengua extranjera puede mantener una relación directa de forma que los informantes que aporten más palabras en español reproduzcan ese modelo en lengua extranjera.

La lectura contribuye a mejorar el conocimiento receptivo del léxico y, a su vez, ese vocabulario pasivo ayuda al desarrollo de la comprensión lectora. Se da por sentado que los informantes que leen más tendrán menos dificultades en reconocer las palabras especializadas y asociarlas con el centro de interés adecuado. En cambio, en el léxico activo esa implicación no será tan plausible ya que la lectura si bien contribuye a incluir nuevas palabras en el repertorio léxico del individuo no conlleva necesariamente al uso de la palabra ni a incorporarlo al vocabulario activo.

La clase social puede influir de forma indirecta en el léxico a través de otras variables como pueden ser el conocimiento de otros idiomas, la realización de viajes y los hábitos de lectura, entre otros factores. No obstante, no se considera un factor relevante ya que la educación ha servido para superar las barreras sociales que llevan a estudiantes a presentar un mayor o menor dominio léxico. En este sentido, Bernstein (1971-1975) analiza el comportamiento lingüístico de las clases sociales y propone la teoría de los códigos lingüísticos (restringido y elaborado). El código restringido es utilizado por la clase trabajadora y depende del contexto. En cambio, el código elaborado es universalista y es empleado por la clase dirigente. En el ámbito educativo, los modelos tienden a reproducirse de forma que provoca el fracaso escolar de los estudiantes de estratos socioeconómicos bajos. De su teoría se deduce

²²¹ Para una revisión teórica de los factores que influyen en el aprendizaje de una lengua puede consultarse el apartado 2.7.

que la mayoría de estudiantes de clase baja no llegan a finalizar los estudios de postgraduado provocando que la clase social no sea tan relevante en hablantes con un alto nivel de instrucción.

En lengua extranjera, la incidencia de los factores puede verse multiplicada debido a que el grado de formación específica en inglés o francés difiere considerablemente de un estudiante a otro. Mientras que todos los informantes han realizado sus estudios universitarios en su totalidad o en parte en lengua española, solo unos cuantos lo han hecho en lengua extranjera. El nivel exigido en el máster es un B1 del Marco Común Europeo de Referencia para las Lenguas que equivale al nivel que adquieren los estudiantes al finalizar la etapa de Bachillerato. Se produce un agravio comparativo entre los sujetos que prácticamente no han continuado con estudios en lengua extranjera desde la etapa preuniversitaria y los informantes que han cursado estudios especializados en lengua extranjera, han realizado estancias en el extranjero, están motivados para aprender otras lenguas y, en definitiva, han continuado con su formación lingüística. En este sentido, la especialidad del máster agrupa a los informantes según los estudios previos ya que el sistema de matriculación establece cuáles son los títulos universitarios desde los cuales se puede acceder a cada especialidad.

En el capítulo anterior, se ha realizado un estudio comparativo general que incluye algunas variables de corte sociolingüístico en comparaciones intramuestrales, como el género, o intermuestrales en las que se analiza el nivel de instrucción o el lugar de residencia de los informantes. Asimismo, se han analizado los vocablos aportados por cada sociolecto para extraer conclusiones sobre la mayor o menor importancia de ciertas unidades léxicas en los hombres o en las mujeres. Desde el punto de vista de la estadística, el sexo no se ha considerado significativo por lo que no forma parte de las hipótesis que se pretenden corroborar ni de los objetivos que se pretenden alcanzar.

En cuanto al léxico pasivo, se presupone que la incidencia de las variables será menor ya que en esa prueba el informante no tiene que producir ninguna palabra técnica sino simplemente identificarla. Se podría decir incluso que algunas de las unidades léxicas que los informantes son capaces de asociar adecuadamente con el

centro de interés no forman parte de su léxico activo. Terrell (1986) distingue dos procesos de transferencia del léxico pasivo al léxico activo: el de ligazón (*binding*) y el de acceso a la palabra (*accessing*).

At first, a new form is recognized as «something familiar», but the meaning (or more accurately, meanings) may not be recalled. Later meaning is slowly recognized. Eventually the form is recognized and interpreted within the phonological stream. Binding is complete when the form evokes the meaning without delay and the form finally «sounds like what it means» (Terrell, 1986: 214).

Cuando un hablante comienza con el aprendizaje de una palabra es capaz de comprenderla pero difícilmente puede escribirla o decirla o lo hace tras una reflexión previa. En cambio, a medida que entra en contacto con ese vocablo se vuelve más accesible de manera que pasa a formar parte del vocabulario activo. En este sentido, se puede afirmar que el léxico disponible forma parte del vocabulario activo mientras que el léxico especializado del vocabulario pasivo y, solo en ocasiones, también del vocabulario activo. De acuerdo con esa hipótesis, las variables sociolingüísticas predisponen en mayor medida los resultados de la prueba de léxico disponible que los del ejercicio de identificación de tecnicismos.

5.3. Análisis descriptivo univariante y presentación de las variables

Según la hipótesis de estudio, se han llevado a cabo diferentes tipos de análisis que requieren variables experimentales de diversa índole: independientes, dependientes e intervinientes. Las variables independientes representan factores, cualidades o características aptas para influir en mayor o menor medida en otras variables. En cambio, las variables dependientes se ven alteradas por los valores de otras variables y, por tanto, su presencia en la realidad de un estudio está condicionada por otras variables. Las variables independientes se representan en el eje de abscisas (X) y las variables dependientes en el eje de coordenadas (Y).

En los análisis de correlaciones influyen tanto en las variables dependientes como independientes algunas variables denominadas extrañas o intervinientes. Estas variables pueden ser ajenas al objeto de estudio y se simbolizan con la letra Z. El atributo o cualidad al que representan estas variables suele ser complejo de delimitar y cuantificar pero su presencia afecta en mayor o menor grado a otras variables.

Siguiendo el ejemplo anterior, la especialidad no es el único factor que influye en el léxico virtual y especializado sino que también intervienen otros como el cansancio y el interés de los alumnos al realizar la prueba. Esos factores, de dificultosa medición, se consideran variables intervinientes en la hipótesis trazada y, consecuentemente, influyentes en las variables expuestas.

El primer paso para realizar una estimación óptima de los resultados pasa por un estudio de sus variables de forma individual que comienza con una presentación de las variables lingüísticas o dependientes por un lado, y de las variables independientes relacionadas con aspectos socioculturales y con el aprendizaje de una lengua extranjera, por otro lado. A través del análisis univariante, se pretende conocer la distribución y dispersión de cada variable de forma aislada. Esta fase es indispensable como aproximación a los análisis bivariantes y multivariantes presentados ulteriormente. Cada uno de estos factores se denomina variable escalar o univariante y el conjunto de las variables dan lugar a una variable vectorial o multivariante.

Para ello, se codificaron todas las variables de forma numérica y se categorizaron según su naturaleza (ordinales, nominales y de escala). A continuación, se realizaron análisis descriptivos que se presentan de dos formas diferentes: mediante diagramas de sectores y barras y a través de tablas con datos exploratorios representados en diagramas de cajas que permiten observar de forma gráfica el análisis de las propiedades de las variables²²².

Por un lado, el diagrama de sectores se ha utilizado para comparar las diversas características que presenta una variable mediante sectores. Por ejemplo, en el caso de la prueba en lengua extranjera, el círculo se ha dividido en dos porciones: una que hace referencia a la lengua inglesa y otra a la lengua francesa. El ángulo central de cada sector es proporcional a la frecuencia absoluta de la lengua y se calcula mediante una regla de tres. Si la muestra es igual a 360 °, la frecuencia de esa cualidad será igual a x .

$$x = \frac{360^\circ \cdot fi}{N}$$

²²² Se han facilitado las descripciones precisas de cada uno de los descriptores para ayudar en la comprensión del capítulo a lectores profanos.

El diagrama de barra o bastones ha sido empleado en los supuestos en que la cualidad observada presenta numerosas opciones, como es el caso de la ocupación del padre y de la madre. Cada barra colocada en el eje de abscisas (X) representa una característica de la variable que muestra su cuantificación en el eje de ordenadas (Y). De esta forma, los valores referentes a la frecuencia absoluta o relativa de cada barra corresponden a la medida expresada en el eje de ordenadas.

Por otro lado, los análisis exploratorios permiten disponer de herramientas de medición más precisas donde los resultados se representan mediante los siguientes parámetros:

- Medidas de tendencia central que muestran mediante diferentes fórmulas la localización de los datos. Concretamente se analizará la media, la mediana y la media recortada al 5 %. La media, también conocida como media aritmética o promedio estándar de los resultados de una muestra (\bar{x}), se calcula sumando todos los valores de un conjunto ($\sum x$) y dividiéndolo por el número de casos o, en este caso, informantes ($\bar{x} = \frac{\sum x}{N}$).
- La mediana calcula el valor central de la variable dejando el mismo número de casos por encima y por debajo. En el caso de las lenguas que habla un informante, si las posibilidades son 1, 2, 3, 4 y 5, la mediana será el número 3. En cambio, si el número de posibilidades fuese par, 1, 2, 3 y 4 la mediana será igual al valor situado entre 2 y 3, es decir, 2,5. La fórmula para calcular la mediana o K , entendido como el valor de la posición de la puntuación en la distribución es la siguiente: $K = \frac{N+1}{2}$

En cambio, la expresión matemática es diferente en los casos en que las distribuciones se encuentran agrupadas en intervalos, siendo la que sigue²²³:

$$Md = Lmd + \left(\frac{\frac{1}{2}N - acum\ fmd}{fmd} \right) \cdot W$$

²²³ Fórmula extraída de García Ferrando (1989: 90). Lmd se refiere al límite inferior del intervalo; N al número total de la muestra, $acum\ fmd$ es la frecuencia acumulada por debajo de la frecuencia del intervalo que presenta la media y W es la distancia de la categoría que contiene la mediana.

- La media recortada al 5 % es un tipo de promedio aritmético que se calcula eliminando el 5 % de los datos más bajos y más altos y, posteriormente, calculando la media con los valores restantes.
- a) Medidas de dispersión que presentan el grado de variabilidad de los valores con respecto a la media por medio de la varianza, la desviación típica, los errores típicos, el mínimo, el máximo, el rango y el rango intercuartílico.
- La varianza (σ^2) de una variable muestra la esperanza del cuadrado de la desviación de esa variable en torno a su promedio.

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

- La desviación típica o desviación estándar (σ o s) es una medida de dispersión alternativa y corresponde a la raíz cuadrada positiva de la varianza y muestra la variación esperada respecto a la media.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

- El rango o recorrido interarticular (R) es la diferencia entre el valor máximo y mínimo de la muestra, es decir, entre el mayor y menor dato de la variable.
- La amplitud intercuartil o rango intercuartílico representa la diferencia entre el primer y tercer cuartil de una media. En general, cuanto menores sean los datos resultantes de las medidas de dispersión mayor será la homogeneidad de la muestra.
- b) Las medidas de la forma de la distribución muestran si los valores se separan o aglomeran según su representación gráfica. Estas medidas describen la manera en que los datos tienden a reunirse de acuerdo con la frecuencia con que se hallen dentro de la información. A través de la asimetría, la curtosis y sus errores típicos es posible conocer las características de la distribución sin tener que recurrir a su gráfico.
- La asimetría refleja si los datos se distribuyen de forma uniforme alrededor del punto central, es decir, de la media. La asimetría presenta tres estados

diferentes según la distribución de los datos con respecto al eje de asimetría: positiva si la mayoría de los valores superan la media aritmética, negativa si esos datos son inferiores a la media y simétrica si los datos se aglomeran de la misma forma por encima y por debajo de la media.

- La curtosis describe el grado en que las proporciones observadas difieren con respecto a la distribución normal representada a través de la conocida campana de Gauss. Al igual que la asimetría, se distinguen tres tipos de curtosis: mesocúrtica, cuando presenta una distribución normal que se considera que oscila entre más y menos 0,5; leptocúrtica, cuando la mayoría de los valores se concentran en la región central de la distribución dando lugar a una curtosis positiva; y platicúrtica en el caso contrario, es decir, cuando la curtosis es negativa y los valores presentan una baja concentración.
- c) El intervalo de confianza estima la probabilidad de que un número aparezca en un rango de valores. El nivel de confianza muestra la probabilidad que el valor se encuentre en el parámetro fijado mientras que el nivel de significación o error aleatorio muestra la posibilidad de fallar en el cálculo estimatorio. El intervalo utilizado en los análisis ha sido del 95 % y, por ende, el nivel de significación es del 5 %. Cuanto mayor sea el rango o intervalo de confianza más preciso es el resultado y viceversa, es decir, cuanto menor sea el intervalo de confianza y mayor la significación mayores son las posibilidades de error.

Los resultados exploratorios se presentan de forma gráfica mediante un diagrama de cajas que, mediante rectángulos, aglomeran los datos entre los cuartiles primero y tercero, es decir entre el 25 % y el 75 % de los datos. El cuadro se encuentra dividido por la mediana que muestra la simetría de los datos ya que cuanto más centrada se encuentre en la caja más simétrica será la variable. Las líneas verticales que sobresalen por los laterales superior e inferior de la caja, conocidas como bigotes, muestran los valores máximos y mínimos de la muestra. Los *outliers* o valores atípicos son identificados fuera de la caja con referencia al número de informante para poder extraer las conclusiones pertinentes.

5.3.1. Variables lingüísticas

5.3.1.1. Índice corregido de descentralización léxica

Esta variable permite conocer el grado de especificidad de las palabras aportadas por un informante con respecto a las unidades léxicas del conjunto de la muestra. El hecho que solo un número reducido de informantes haya mencionado ciertos vocablos puede atender a varios motivos: por tratarse de asociaciones individuales o simplemente poco frecuentes y por ser palabras propias de un lenguaje más especializado.

El índice de descentralización contabiliza todas las palabras con el valor 1 sin ningún tipo de consideración previa sobre la pertinencia más o menos directa del vocablo con el centro de interés. En cambio, Ávila Muñoz y Villena Ponsoda (2010: 185) proponen corregir este índice aplicando valores diferenciados según el tipo de relación de los vocablos con respecto al centro de interés o núcleo prototípico: palabras nucleares ($N1 = 1$), asociaciones colectivas ($N5 = 0,5$), derivaciones ($N25 = 0,25$) y asociaciones individuales ($N0 = 0$). La suma de las palabras previamente ponderadas según esta clasificación da lugar al índice corregido de descentralización léxica²²⁴.

La selección del ICoDD como variable dependiente es fruto de un análisis pormenorizado a nivel estadístico y teórico de las posibles variables que pueden medir la disponibilidad léxica: el número total de palabras aportadas por un informante, el índice corregido de amplitud léxica o lo que es lo mismo la ponderación cualitativa del total de palabras y el propio índice corregido de descentralización léxica. Por una parte, la correlación bilateral del ICoDD con el número de palabras y el índice corregido de amplitud léxica es significativa a nivel 0,01 (bilateral) y el coeficiente de Pearson es superior a 0,9 en ambos casos por lo que sería redundante utilizar estas tres variables lingüísticas. Por otra parte, este índice es capaz de aportar información más precisa sobre el léxico disponible que la mera suma de las palabras.

²²⁴ *Vid.* apartado 2.5.4 para una revisión más detallada de las características de la variable.

La Tabla 5.1 muestra que los resultados del caudal léxico en ambas lenguas son similares aunque los resultados son superiores en español que en lengua extranjera. La media en español es de 0,48 mientras que en lengua extranjera es 0,47. El error típico es bajo en ambos casos, en lengua extranjera 0,009 y 0,006 en español. En cuanto a la concentración de los datos, en español son más compactos ya que la desviación típica es de 0,0898 y en lengua extranjera de 0,1223. La curtosis en sendos análisis es negativa pero al ser inferior a -0,5 la concentración de los datos se considera normal o mesocúrtica.

No obstante, en el cálculo de la descentralización influye el número de palabras en cada centro y el tamaño de la muestra. Por un lado, si los hablantes han aportado pocas palabras es probable que haya ciertas palabras que solo hayan sido aportadas por varios informantes y se consideren poco compatibles. Por otro lado, el número de informantes influye de forma que cuanto menor sea la muestra mayor puede ser el número de palabras descentralizadas. A tenor de lo dicho, se pueden comparar, *grosso modo*, las dos muestras pero los resultados deben comprenderse de forma separada ya que en lengua extranjera los informantes han aportado un menor número de palabras y se han dividido según el idioma.

ICoDD	Media		Desv. típ.		Varianza		Asimetría		Curtosis	
	Estad.	Error T.	Estad.	Estad.	Estad.	Error T.	Estad.	Error T.	Estad.	Error T.
En lengua extranjera	0,479275	0,006864	0,089751	0,008	0,008	0,186	-0,136	0,369		
En español	0,466631	0,009351	0,122274	0,015	0,209	0,186	-0,454	0,369		

Tabla 5.1. Análisis exploratorios sobre el ICoDD en lengua extranjera y en lengua española

5.3.1.2. Índice corregido de identificación de tecnicismos (ICIT)

5.3.1.2.1. Índice corregido de identificación de tecnicismos en lengua extranjera

En el estudio de Ávila Muñoz y Villena Ponsoda (2010: 217), los resultados de la prueba de léxico especializado se mostraron a través del Pident definido como la capacidad individual de identificación de piezas léxicas especializadas. Este índice se obtiene de la sumatoria de todas las respuestas correctas de los informantes sin tener en cuenta el azar implícito en la prueba. Para corregir el índice de identificación de tecnicismos, se ha restado a la suma de las respuestas correctas aquellas respuestas

incorrectas. A las primeras se les da el valor de 0,56 y a las segundas de 0,7 calculado según la distribución de la probabilidad²²⁵. Las respuestas no contestadas no puntúan. Los resultados del Pident y el ICIT tienen forma de escala que va desde 0 hasta 10 puntos, en el primer caso, y desde valores inferiores a 0 hasta 10, en el segundo caso. Los datos se han distribuido en cuatro rangos en el caso del Pident y en seis en la variable ICIT (dos más correspondientes a valores inferiores a 0)²²⁶.

Pident en LE por rangos	Casos	ICIT en LE por rangos	Casos
Rangos		<-025 %	9
		-0,25-0 %	47
0-25 %	59	0-25 %	72
25-50 %	84	25-50 %	30
50-75 %	24	50-75 %	10
75-100 %	4	75-100 %	3

Tabla 5.2. Comparativa del Pident y el ICIT en lengua extranjera por rangos

En la columna referida al Pident la mayoría de los resultados se concentran en el primer y segundo rango, es decir, entre el 0 hasta el 25 % y entre el 25 y 50 % con 59 y 84 informantes respectivamente. En cambio, en el ICIT los datos se aglutinan en el segundo, tercer y cuarto rango con 47, 72 y 30 informantes. Si esta prueba fuese un examen habrían aprobado según el Pident 28 informantes ya que superan el umbral de 5 puntos, de los cuales 24 informantes se sitúan en el nivel 3 con una nota de 5 a 7,5 y tan solo 4 en el último escalón, con una calificación de 7,5 a 10. En contraposición, según los resultados del ICIT solo hubieran superado la prueba 13 informantes, 10 con una calificación entre 5 y 7,5 y solo 3 con una nota superior a 7,5.

5.3.1.2.2. Índice corregido de identificación de tecnicismos en lengua española

En el caso del léxico especializado en lengua española, la media de los resultados en el Pident y el ICIT son superiores ya que tan solo 60 informantes se sitúan por debajo del 50 % según el cálculo del Pident y 100 según el ICIT en contraposición con los 143 y 158 informantes que se situaban bajo ese umbral en

²²⁵ Vid. apartado 2.5.5 para más información sobre el cálculo probabilístico de el ICIT.

²²⁶ Aunque los datos presenten la exploración de dos variables se ha incluido en este apartado de presentación de variables para explicar el cálculo de la variable y compararla con otro índice.

lengua extranjera. El grueso de las respuestas en el Pident se localiza en orden decreciente como sigue: el primero es el rango correspondiente al 25-50 %, le siguen los resultados entre el 50-75 %, el 0-25 % y el 75-100 %. En el ICIT se constata que tan solo siete informantes han realizado la prueba con valores inferiores a 0.

Pident en español por rangos	Casos	ICIT en español por rangos	Casos
Rangos		<-025 %	7
		-0,25-0 %	0
0-25 %	6	0-25 %	33
25-50 %	54	25-50 %	60
50-75 %	81	50-75 %	53
75-100 %	30	75-100 %	18

Tabla 5.3. Comparativa del Pident y el ICIT en español por rangos

5.3.2. Variables sociodemográficas y reticulares

5.3.2.1. Edad

Al tratarse de estudiantes de posgrado, el alumnado ha debido realizar previamente una licenciatura o diplomatura²²⁷ y, por tanto, su edad oscila entre los 22 y los 49 años. Los datos descriptivos de la Tabla 5.4 muestran una media de edad de 28,73 años y una concentración de los informantes entre los 22 y 30 años. Las respuestas de los informantes se aglutinan en el primer cuartil produciéndose un claro descenso en el número de encuestados a partir de los 30 años, lo cual supone que los estudiantes realizan el máster inmediatamente después de finalizar sus estudios universitarios.

	Análisis descriptivos	Estadístico	Error típ.
Edad	Media	28,731	0,46031
	Intervalo de confianza para la media al 95 %	Límite inferior	27,8223
		Límite superior	29,6397
	Media recortada al 5 %	28,1404	
	Mediana	28	

²²⁷ Los datos fueron recabados en el año 2011 por lo que ningún estudiante finalizó los estudios de grado. La implantación del título de grado en España está regulada por el Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales (BOE 30-10-07) y las modificaciones introducidas en el Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado (BOE 29/01/2011) y el Real Decreto 861/2010, de 2 de julio, por el que se modifica el Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales (BOE 03/07/2010).

Varianza	36,233	
Desv. típ.	6,01939	
Mínimo	22	
Máximo	49	
Rango	27	
Amplitud intercuartil	7	
Asimetría	1,348	0,186
Curtosis	1,581	0,369

Tabla 5.4. Análisis exploratorios de la edad de los informantes

Los resultados exploratorios del Gráfico 5.1 muestran una distribución normal cuya mediana se desplaza ligeramente por encima de la media. A partir de 43 años los estudiantes se consideran casos extremos²²⁸ ya que la mayoría de los sujetos son recién titulados. A diferencia de otros estudios de postgrado, este máster conduce a una profesión específica que exige esta certificación como requisito indispensable por lo que la mayoría de los estudiantes no han trabajado anteriormente sino que lo realizan como salida al mundo laboral.

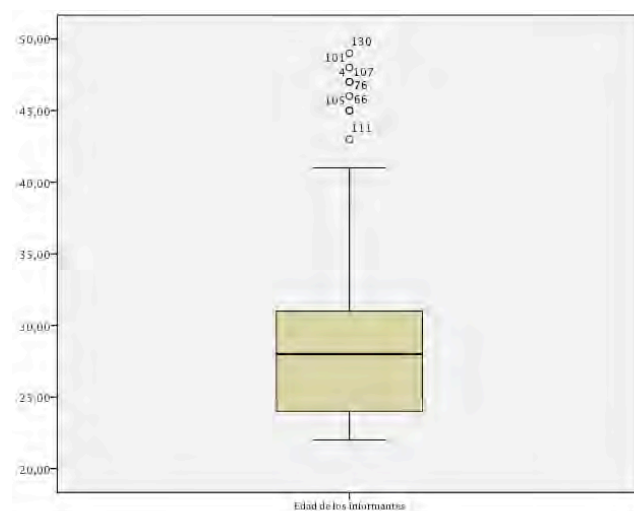


Gráfico 5.1. Gráfico de tallos y hojas de la edad de los informantes

Esta homogeneidad conlleva que la edad no tenga una incidencia relevante en las variables dependientes. En los estudios herederos del Proyecto Panhispánico no es frecuente encontrar la edad como un factor condicionante de la disponibilidad léxica ya que la muestra es más uniforme aún: la mayoría de los informantes tiene

²²⁸ Vid. los informantes marcados en el gráfico de tallos y hojas: 4, 66, 76, 101, 105, 107, 111, 130.

entre 17 y 18 años, salvo casos excepcionales en los que pueden ser menores y alumnos repetidores. Sin embargo, la edad se revela como una variable destacada en estudios comparativos con muestras diferenciadas, como es el caso de los estudiantes preuniversitarios y los del estudio actual. En ese supuesto, la edad es un factor oculto que viene condicionado realmente por el nivel de estudios de los informantes: preuniversitarios y postgraduados.

5.3.2.2. Estudios

En el cuestionario sociológico, se han incluido cuatro preguntas relativas a los estudios del informante y de sus padres²²⁹. En primer lugar, el estudiante proporciona datos sobre su formación académica: nombre de la carrera universitaria, tipo de título universitario (distinguiendo entre diplomatura, licenciatura o ambos); otros estudios del informante (como puede ser la realización de otro posgrado) y años concretos de estudio. Para la codificación de los años de estudios, además de incluir una variable con el dato exacto se creó una variable que divide la muestra en tres grupos: de 15 a 20 años de estudios, de 21 a 26 y más de 26 años.

El Gráfico 5.2 muestra un ejemplo de la distribución por tipos de estudios donde se constata que tres cuartas partes de los informantes son licenciados, seguidos de los diplomados y de una minoría que realizó una diplomatura y licenciatura.

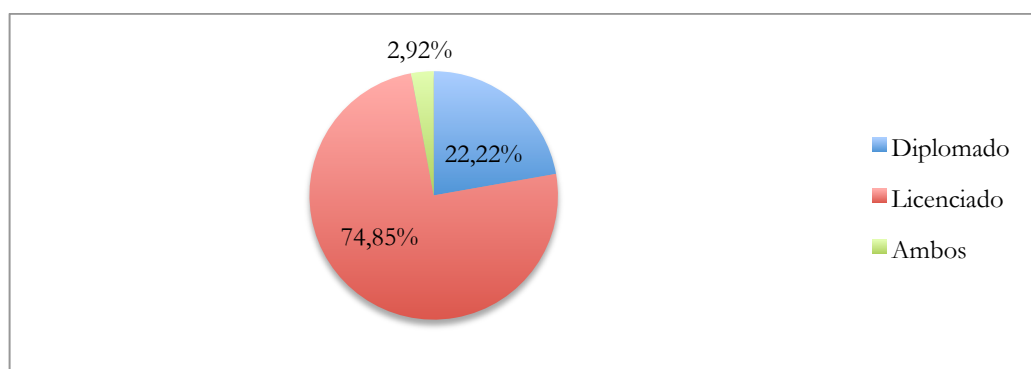


Gráfico 5.2. Diagrama de sectores de los estudios del informante

En cuanto a la formación de los padres, se solicitó información sobre los años de estudios y el título máximo alcanzado. El primer apartado era desconocido por casi un 30 % de los informantes razón por la cual la variable seleccionada para el

²²⁹ Vid. preguntas 9 a la 12 del cuestionario en el Anexo IV.

análisis estadístico es el nivel de estudios alcanzado ya que a su vez da cuenta de los años académicos cursados por los padres del informante. Los estudios de los padres es un componente que forma parte de la variable del nivel sociocultural incluida en la mayoría las investigaciones herederas del Proyecto Panhispánico ya que la repercusión sobre el léxico disponible suele ser significativa:

Los listados de léxico disponible aportan también interesantes datos sobre la influencia de las variables sociales en el dominio léxico de los informantes. En general, podemos decir que la variable que más influencia parece tener sobre el dominio léxico es el nivel cultural de la familia, más que la variable sexo, o la pertenencia a un ambiente urbano o rural. La comparación entre las distintas variantes de una variable nos informa de la diferente integración de algunas palabras en la sociedad (Bartol Hernández, 2010: 103).

El Gráfico 5.3 muestra el nivel de estudios alcanzados por el padre y la madre. Como se observa, el sexo masculino ha logrado un nivel de estudios superior que el femenino en líneas generales. Los padres con máster o doctorado alcanzan un 2,35 % y con una licenciatura 18,82 % mientras que las madres con doctorado apenas llegan al 0,59 % y al 12,35 % con licenciatura.

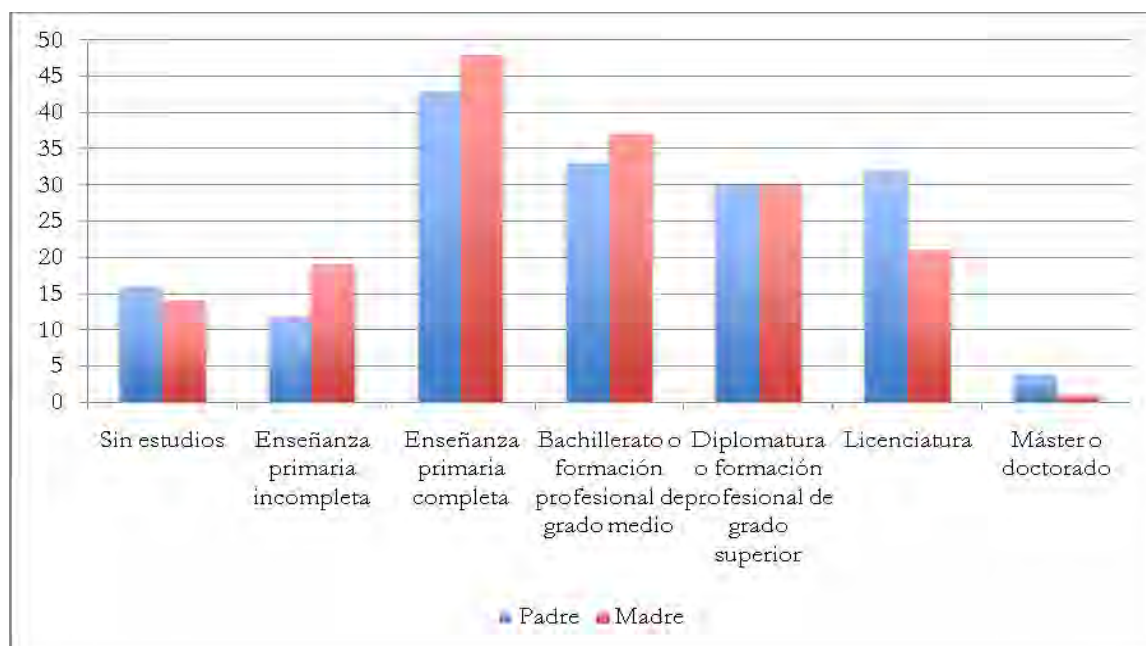


Gráfico 5.3. Diagrama de barras de los estudios de los padres

5.3.2.3. Ocupación de los padres e ingresos

La codificación de las profesiones se realizó siguiendo las pautas marcadas en el Real Decreto 1591/2010, de 26 de noviembre, por el que se aprueba la Clasificación Nacional de Ocupaciones 2011. Esta supone la última revisión en el

momento de la encuesta²³⁰ de la estructura de la Clasificación Nacional de Ocupaciones (CNO-94) con objeto de incorporar nuevas ocupaciones sobre todo en el ámbito de las Tecnologías de la Información y las Comunicaciones y del Medio Ambiente.

Esta nueva clasificación, vigente desde el 1 de enero de 2011, recoge nueve categorías de ocupaciones subdivididas en diferentes tipos de profesiones para dar una información más concreta del tipo de profesión. La ordenación parte desde las ocupaciones de directores y gerentes a ocupaciones elementales. Los enunciados de esas categorías, junto con dos categorías que complementan la ocupación de los familiares de los hablantes (ama de casa y desempleado), son los siguientes:

1. Directores y gerentes.
2. Técnicos y profesionales científicos e intelectuales.
3. Técnicos y profesionales de apoyo.
4. Empleados contables, administrativos y otros empleados de oficina.
5. Trabajadores de los servicios de restauración, personales, protección y vendedores.
6. Trabajadores cualificados en el sector agrícola, ganadero, forestal y pesquero.
7. Artesanos y trabajadores cualificados de las industrias manufactureras y la construcción (excepto operadores de instalaciones y maquinaria).
8. Operadores de instalaciones y maquinaria, y montadores.
9. Ocupaciones elementales.
10. Ama de casa.
11. Desempleado.

²³⁰ La encuesta se realizó entre enero y abril de 2011.

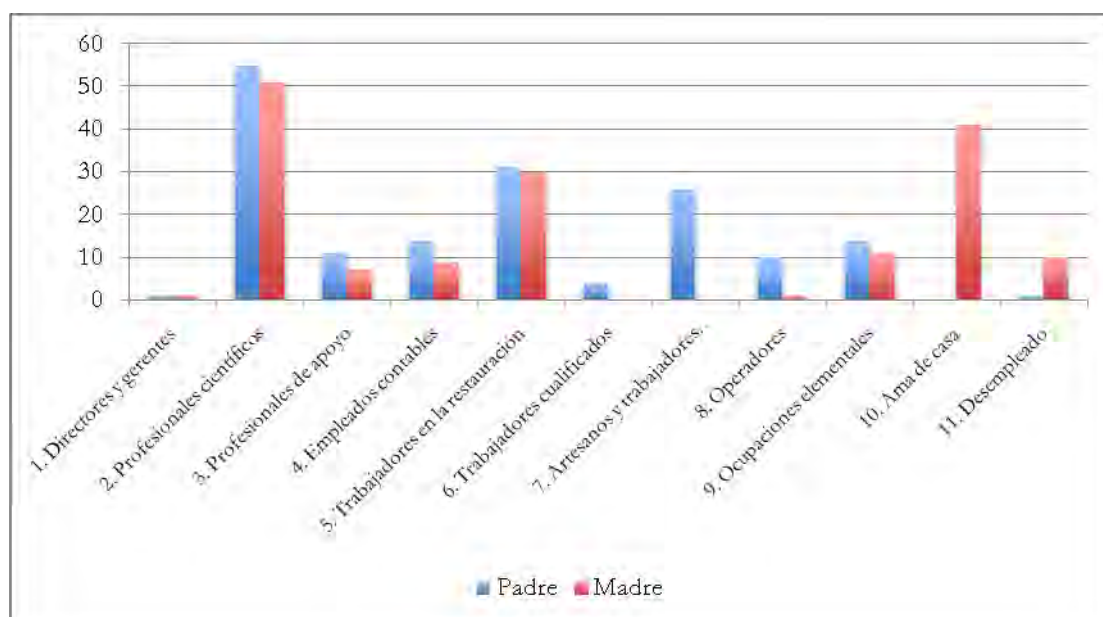


Gráfico 5.4. Diagrama de barras de la ocupación de los padres

Las ocupaciones de los padres también reflejan una ligera supremacía en el puesto ocupado por los hombres en comparación con el de las mujeres, lo que concuerda con la formación mostrada en el apartado anterior. Hay un 25,47 % de madres desempleadas y un 6,21 % de amas de casa mientras que solo hay un 8,38 % de hombres desempleados y ninguno se dedica en exclusividad a las labores domésticas. En las ocupaciones más prestigiosas, el porcentaje de hombres es superior al de mujeres: un 32,68 % frente a un 31,68 % son técnicos y profesionales científicos e intelectuales.

En el apartado de ingresos familiares, se solicitó incluir la renta familiar de los padres para poder así completar la clase social de la que provienen. Tan solo una minoría de informantes está emancipada y se corresponde con los informantes con una edad superior a la media. Solo hay un estudiante cuya renta familiar es inferior a 400 euros al mes y nueve informantes con unos ingresos mensuales entre 400 y 799 euros.

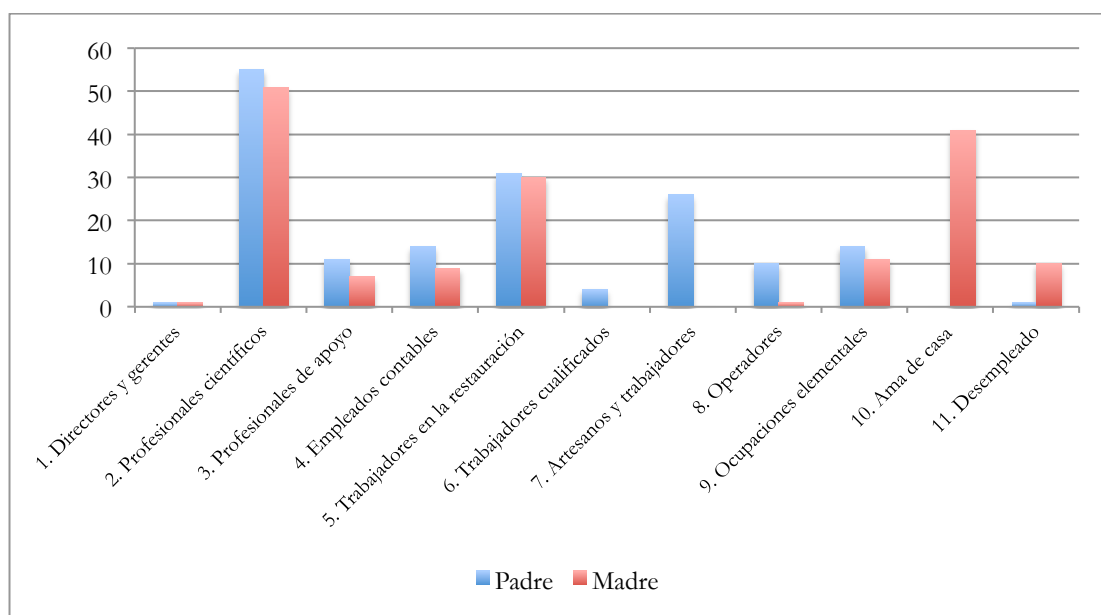


Gráfico 5.5. Gráfico de sectores sobre los ingresos familiares

El precio del crédito del máster en el curso 2010/2011 era de 27,60 euros y el número de créditos es 60 de forma que cada estudiante debió pagar 1656 más gastos de seguro escolar y expedición del carné universitario, salvo aquellos con exenciones de pago o becados, con lo que dificulta el acceso de aquellos estudiantes con rentas bajas. Cabe destacar que los precios públicos para las tasas y prestación de servicios universitarios de las Universidades Públicas se fijan y publican anualmente por la Junta de Andalucía²³¹.

5.3.3. Variables relacionadas con el dominio de una lengua

Tal y como se ha referido en la fundamentación teórica²³², en el proceso de adquisición de una lengua extranjera intervienen diferentes factores, tales como la edad, el tipo de metodología empleada y el contexto. Para conocer en qué medida influye el contexto de aprendizaje de la lengua extranjera en el vocabulario activo y pasivo de los informantes en lengua extranjera, se incluyen consultas en el cuestionario sociológico que hacen referencia a la edad de aprendizaje de una lengua,

²³¹ Para el curso 2010/2011 el precio se fijó según el Decreto 329/2010, de 13 de julio, por el que se autoriza la implantación de enseñanzas universitarias de Grado, Máster y Doctorado, se actualiza el Catálogo de Enseñanzas Universitarias conducentes a la expedición por las Universidades Públicas de Andalucía de títulos oficiales, así como la estructura de los centros que las imparten y se fijan los precios públicos y tasas a satisfacer por la prestación de servicios académicos y administrativos universitarios para el curso 2010/2011. Boletín número 139 de 16/07/2010.

²³² *Vid.* apartado 2.7.

pudiendo ser codificados los resultados según la edad exacta o por tramos de edad; el lugar de aprendizaje (colegio, escuela de idiomas y academia privada, entre otros); el tipo de centro educativo ha cursado los estudios de primaria y secundaria (público, privado o concertado), etc.

En cuanto al aprendizaje en el ámbito académico universitario, el informante ha contestado preguntas sobre la realización de asignaturas en lengua extranjera durante la carrera y la especialidad del máster que cursa. Asimismo, se espera que esas variables y otras relacionadas con los hábitos de lectura y la exposición a los medios de comunicación tengan un impacto en la lengua materna aunque en menor medida que en lengua extranjera.

A continuación, se comentan cinco agrupaciones de variables que engloban factores que inciden en el vocabulario en español y en lengua extranjera. La primera se refiere a las especialidades que se ofertan en el máster y los estudios universitarios desde los cuales se puede acceder. La segunda recoge el número de idiomas que conoce el sujeto y el nivel de competencia lingüística que ha adquirido según el Marco Común Europeo de Referencia para las Lenguas. El tercer grupo de variables está relacionado con el uso de la lengua en diferentes ámbitos con indicación del tipo de lengua (lengua española, lengua extranjera de la prueba u otra lengua). El cuarto hace referencia a la exposición a los medios de comunicación y a los hábitos de lectura y recoge los datos de los libros que el informante lee en español. El quinto y último grupo recoge las variables que muestran la motivación del informante para impartir clases bilingües y proporciona los datos de una de las preguntas.

5.3.3.1. Especialidad del máster

La Orden ECI/3858/2007 de 27 de diciembre por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habilitan para el ejercicio de las profesiones de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas, establece además sus objetivos y planificación académica. Por ello, tal y como indica el propio nombre del máster, las enseñanzas se estructuran conforme a las materias y ámbitos docentes de la Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y

Enseñanza de Idiomas. El máster incluye un módulo específico para cada especialidad compuesto por 24 créditos que engloba tanto contenidos curriculares de las materias correspondientes como conocimientos didácticos²³³.

Para establecer la pertinencia entre titulaciones y la especialidad del máster, la Comisión de Distrito Único Andaluz revisa anualmente la relación existente entre ambas. De esta forma, se realiza una clasificación de las titulaciones según sea de preferencia alta, media o baja para determinar la admisión del alumnado en cada especialidad, posicionándose en primer lugar el alumnado cuya titulación sea de preferencia alta y en último aquel que posea una titulación de preferencia baja.

Cabe destacar que si un alumno solicita matricularse en una especialidad diferente a la de acceso directo, debe realizar al inicio del curso una prueba relacionada con la especialidad. Ese examen está relacionado a su vez con el temario aprobado para el procedimiento selectivo para el ingreso en los cuerpos de profesores de enseñanza secundaria, profesores técnicos de formación profesional, profesores de escuelas oficiales de idiomas, profesores de música y artes escénicas, profesores de artes plásticas y diseño y maestros de taller de artes plásticas y diseño.

Teniendo en cuenta la relación directa que se establece entre la titulación del informante y la especialidad del máster, en esta investigación la variable denominada «Especialidad» hace referencia no solo a la especialidad sino también a la formación previa del alumnado. En cuanto a la organización de esas especialidades se ha tomado como referencia el Decreto 327/2010, de 13 de julio, por el que se aprueba el Reglamento Orgánico de los Institutos de Educación Secundaria. En la normativa citada, se establecen las áreas de competencias y los conocimientos presentes en cada una de ellas así como la organización de las diferentes asignaturas. El artículo 84 distingue las siguientes áreas:

- Área social-lingüística: cuyas competencias principales son la competencia en comunicación lingüística (tanto en lengua española como en lengua extranjera) y la competencia social y ciudadana.

²³³ En el apartado 2.4 se detalla el plan de estudios del máster.

- Área científico-tecnológica: comprende la competencia de razonamiento matemático, competencia en el conocimiento y la interacción con el mundo físico y natural y la competencia digital y tratamiento de la información.
- Área artística: centrada en desarrollar la competencia cultural y artística.
- Área de formación profesional: se centra en las competencias profesionales presentes en las enseñanzas de formación profesional existentes en el centro.

En la siguiente tabla, se expone el tipo de estudios a través de los cuales los informantes han accedido a cada especialidad junto con el área al que corresponde cada una de ellas. Además, se ha incluido el número exacto de informantes que han accedido con cada titulación según la especialidad.

Área	Especialidad	Titulación	Nº informantes
Social-lingüística	Lengua Extranjera: Inglés	Ldo. en Filología Inglesa	13
		Ldo. en Traducción e Interpretación	11
		Ldo. en Filología Hispánica	1
	Lengua Extranjera: Francés	Ldo. en Filología Inglesa	1
		Ldo. en Filología Francesa	2
		Ldo. en Traducción e Interpretación	7
		Ldo. en Filología Hispánica	3
	Lengua y Literatura. Latín y Griego	Ldo. en Filología Francesa	2
		Ldo. en Traducción e Interpretación	4
		Ldo. en Filología Hispánica	5
		Ldo. en Filología Clásica	2
		Ldo. en Ciencias de la Comunicación	1
		Ldo. en Periodismo	1
		Ldo. en Comunicación Audiovisual	1
		Ldo. en Publicidad y Relaciones Públicas	3
	Geografía e Historia	Ldo. en Ciencias de la Comunicación	1
		Ldo. en Geografía	3
		Ldo. en Historia o Historia del Arte	10
		Ldo. en Filosofía	3
Científico-tecnológica	Tecnología, Informática y Procesos Industriales (TIPI)	Ldo. en Arquitectura	3
		Ldo. en Informática	3
		Ingeniero químico	1
		Ldo. en Ingeniería de Telecomunicaciones	2
		Ldo. en Química	1
		Diplomado en Informática (de Sistemas, de Gestión)	4
		Diplomado en Ingeniero Técnico de Telecomunicaciones	10
		Diplomado en Ingeniería Técnica Industrial	8

	Matemáticas	Diplomado en Ingeniería Técnica Informática	4
		Diplomado en Arquitectura Técnica	4
		Ldo. en Matemáticas	7
		Ldo. en Arquitectura	1
		Ingeniero químico	1
		Ldo. en Ingeniería Industrial	1
Artística	Dibujo, Imagen y Artes Plásticas (DIAP)	Ldo. en Publicidad y Relaciones Públicas	2
		Ldo. en Bellas Artes	7
		Ldo. en Arquitectura	2
	Música	Ldo. en Arte Dramático	3
		Ldo. en Música	6
Formación Profesional	Orientación Educativa	Ldo. en Publicidad y Relaciones	1
		Ldo. en Psicología	7
		Ldo. en Pedagogía	2
	Formación y Orientación Laboral (FOL)	Ldo. en Psicología	1
		Ldo. en Derecho	3
		Ldo. en Políticas y Sociología	1
		Ldo. en Ciencias del Trabajo	1
		Diplomado en Relaciones Laborales	1
	Hostelería y Turismo	Ldo. en Publicidad y Relaciones	1
		Diplomado en Turismo	9

Tabla 5.5. Tabla que muestra los diferentes estudios universitarios divididos por áreas

5.3.3.2. Conocimiento de otras lenguas extranjeras

Al igual que en investigaciones previas (Carcedo González, 2000c; Samper Hernández, 2011; González Fernández, entre otros), los informantes indicaron en el cuestionario sociolingüístico el número de idiomas que conocen y su nivel de dominio según el Marco Común Europeo de Referencia para las Lenguas. El objetivo era averiguar en qué medida el número de lenguas que habla un individuo influye en su léxico. La selección de esta variable intenta respaldar las teorías que afirman que el aprendizaje de más de un idioma favorece el aprendizaje de otros, mejorando incluso los idiomas ya aprendidos.

El Gráfico 5.1 muestra que más de la mitad de los estudiantes, concretamente el 69,59 %, conoce una o dos lenguas extranjeras. Según la edad de los informantes, la mayoría ha estudiado en el plan de estudios que establece la Ley Orgánica General del Sistema Educativo (LOGSE) de 1990 en la que se implanta la obligatoriedad de una segunda lengua extranjera en Bachillerato. En plan de estudios

explicaría, en cierto modo, que la media de idiomas que conoce un informante sea de 2,0117.

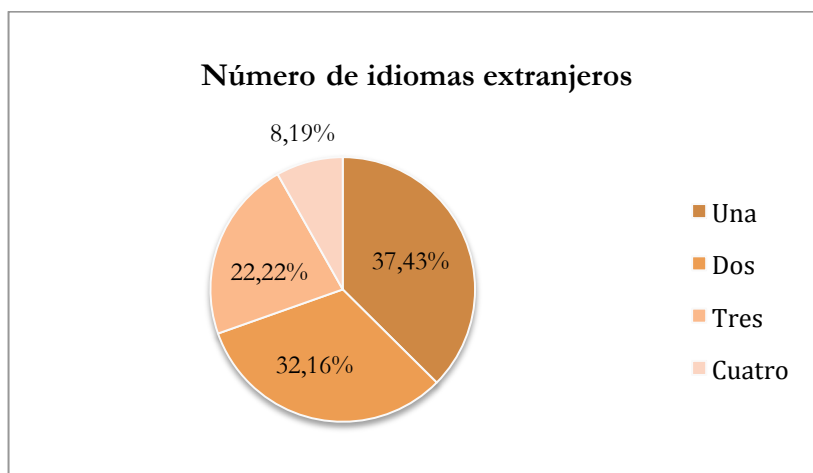


Gráfico 5.6. Diagrama de sectores del número de idiomas extranjeros que conoce el informante

Una proporción considerable de individuos, más de un cuarto, afirma conocer tres o cuatro idiomas por lo probablemente se trate de informantes que han cursado estudios universitarios relacionados con lenguas extranjeras. Dentro del grupo de estudiantes que habla un idioma es necesario establecer diferencias según los niveles alcanzados en cada lengua para conocer si el estudiante ha continuado sus estudios en lengua extranjera o, si por el contrario, la última vez que estudió inglés o francés fue en Bachillerato. En el apartado de análisis bivariantes se analiza con detenimiento la relación entre el léxico y el número y nivel de idioma alcanzado por el hablante.

5.3.3.3. Uso de la lengua extranjera en diferentes contextos

Además de indagar sobre las lenguas que conocen el informante y sus familiares (padres, hermanos y pareja), se solicitó información sobre la lengua vehicular utilizada con las personas de su entorno: en clase, con amigos, con la familia y en otra situación como puede ser en el ámbito laboral. La frecuencia de uso está organizada en forma de escala que oscila entre el uso diario hasta la no utilización de la lengua. Se parte de la hipótesis que cuanto mayor sea el contacto con la lengua extranjera mejores serán los resultados en las pruebas de léxico.

En el ámbito académico, la mayoría de los hablantes nunca usa la lengua extranjera de la prueba, concretamente un 62,57 %, mientras que tan solo un 12,28 %

la usan en clase. Esta distribución corresponde a estudiantes de las especialidades del máster en lengua extranjera y de otros alumnos que se encuentran realizando estudios de idiomas extranjeros al margen del máster.

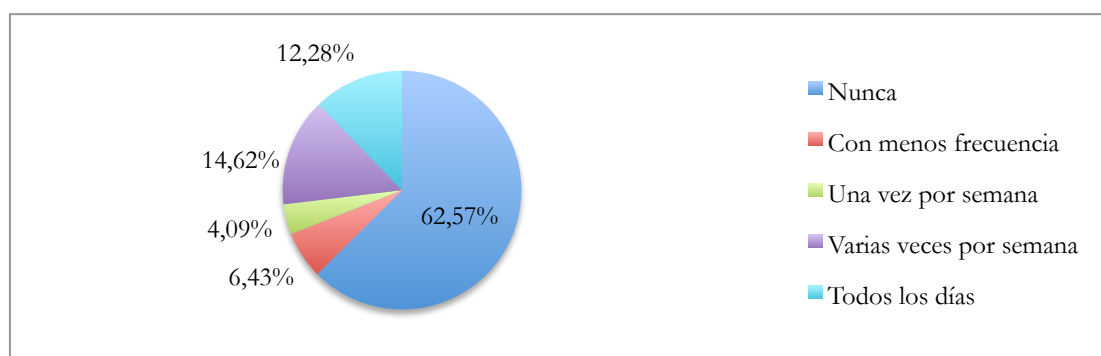


Gráfico 5.7. Gráfico de sectores sobre la frecuencia en el uso de la lengua extranjera de la prueba en clase

Asimismo, se considera clave el uso de la lengua extranjera en el contexto de inmersión a través de las estancias en el extranjero de habla no hispana. Esta variable está correlacionada significativamente con el uso de la lengua extranjera en diferentes contextos ya que implica necesariamente el intercambio lingüístico en un idioma extranjero con otras personas. En el apartado 5.4.3.3 se ha relacionado el tipo de estancias en el extranjero con las variables lingüísticas para comprobar si existe relación entre ellas. Asimismo, esos datos se completan en el análisis de componentes principales del apartado 5.5.1.3 donde se pretende reducir el número de variables que se refieren al uso de la lengua en contexto como puede ser el ámbito profesional, académico, personal, en estancias lingüísticas y en relación con los medios de comunicación.

5.3.3.4. Exposición a los medios de comunicación y libros al año en español

Al igual que en el estudio de Ávila Muñoz y Villena Ponsoda (2010), se han incluido preguntas relacionadas con la exposición a los medios de comunicación. Para facilitar la obtención de los datos, los informantes marcaron con una cruz las horas semanales que dedican a ver la televisión según el tipo de programas (documentales, programas de noticias del corazón, entre otros); a leer periódicos; a buscar información en internet; a conectarse a las redes sociales; a ver películas y a

escuchar música. Los tramos horarios van desde ninguna hora, pasando por menos de dos horas, entre dos y cinco hasta más de cinco horas²³⁴.

La particularidad de los datos viene dada por la especificidad de la lengua utilizada y se ha establecido la distinción entre la lengua española (LE), la lengua del test (LET) y otra lengua extranjera (LE2)²³⁵. En lengua extranjera, la exposición a los medios puede ser un factor que influye positivamente en el aprendizaje y que está relacionado con el uso de la lengua en diferentes contextos. En cambio, en lengua española los hábitos de lectura así como la exposición a programas culturales pueden incidir en el vocabulario activo y pasivo de los informantes pero no está ligado al uso de la lengua en diferentes contextos ya que sería redundante preguntar a los estudiantes si usan el español como lengua vehicular con amigos o familiares. En el Gráfico 5.8, se constata que la mayoría de los estudiantes leen cinco o menos de cinco libros al año y solo un 2,34 % afirma no leer ningún libro.

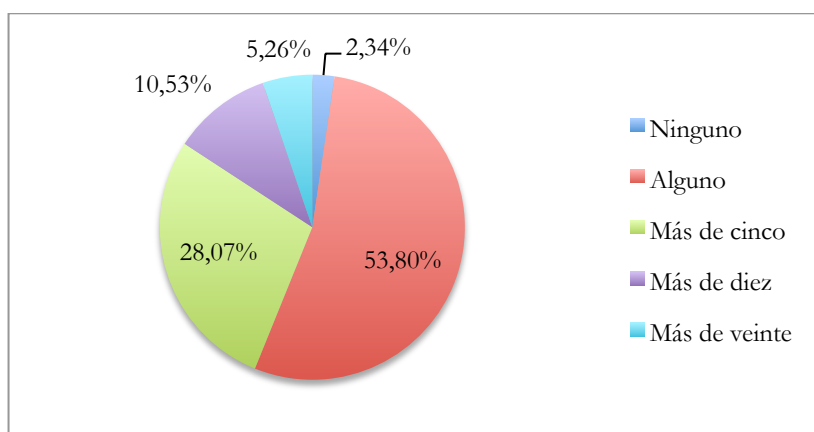


Gráfico 5.8. Gráfico de sectores sobre el número de libros que el informante lee al año

5.3.3.5. Motivación del futuro profesorado

Durante la fase de formación inicial del profesorado, resulta imprescindible conocer cuál es la motivación de los estudiantes para impartir clases en lengua extranjera y su opinión en cuanto a la formación recibida²³⁶. Por un lado, los informantes de las especialidades del máster en lengua extranjera inglés o francés valoraron la idoneidad del uso de la lengua extranjera como lengua vehicular en clase

²³⁴ Vid. pregunta 26 del cuestionario sociolingüístico del Anexo IV.

²³⁵ Para más detalles, véase la pregunta 27 del cuestionario sociolingüístico situado en el Anexo IV.

²³⁶ A tal efecto, se han dedicado las trece últimas preguntas del cuestionario.

frente a la lengua materna. Por otro lado, el alumnado de otras especialidades dio su opinión sobre la necesidad de incluir alguna clase en lengua extranjera para poder prepararse para dar clases bilingües.

En cuanto a la práctica docente, los estudiantes comentaron las razones por las cuales les gustaría o no llevar a cabo las prácticas en centros bilingües, si conocían el programa del Aprendizaje Integrado de Contenidos en Lengua Extranjera (AICLE) y cuál era su sensación al imaginarse impartiendo clases de algún área no lingüística en lengua extranjera.

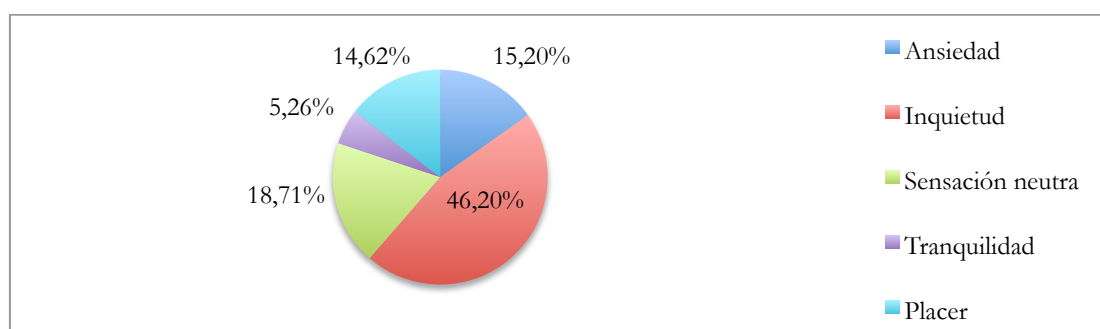


Gráfico 5.9. Diagrama de sectores sobre la sensación del informante al impartir clases en lengua extranjera

La mayoría de los informantes afirma sentir ansiedad o inquietud al imaginarse impartiendo una clase en la que se enseñen los contenidos curriculares en español y en lengua extranjera. Tan solo un 14,62 % afirma que le gustaría impartir clases en lengua extranjera de los cuales más del 90 % corresponde a profesores de inglés y francés. Esta variable engloba otros factores intrínsecos como la seguridad misma del informante sobre su formación, la actitud ante la participación en un programa bilingüe y la difusión de programaciones y actividades bilingües en el máster.

5.3.4. Resumen de las variables

A continuación, se presentan de forma esquemática las principales variables analizadas en este estudio con indicación de la media y la desviación típica de los sociolectos masculino y femenino de forma conjunta y separada. Cabe destacar que se ha aplicado la prueba *t* de Student para comparar la posible diferencia ente el sexo de los informantes pero no ha sido significativa. Siguiendo el orden establecido en el análisis univariante, se sitúan en primer lugar los datos globales de las variables

lingüísticas, seguida de las variables socioculturales y relacionadas con el aprendizaje de una lengua.

Variables lingüísticas (variables dependientes)	Lengua	Total			Hombres			Mujeres		
		Media	Desv.	N	Media	Desv.	N	Media	Desv.	N
Índice corregido de descentralización léxica	Español	0,48	0,47	171	0,47	0,09	54	0,48	0,09	117
	LE	0,47	0,45	171	0,46	0,13	54	0,47	0,12	117
Índice corregido de amplitud léxica	Español	181,64	33,12	171	173,62	34,3	54	185,34	32,03	117
	LE			171	120,6	33,29	54	126,88	32,28	117
Índice corregido de identificación de tecnicismos	Español	5,64	1,82	171	5,93	1,9	54	5,5	1,77	117
	LE	3,08	1,87	171	3,31	1,76	54	2,97	1,92	117
Asociaciones individuales (N0)	Español	1,59	2,05	171	1,78	1,89	54	1,50	2,13	117
	LE	2,87	2,98	171	2,82	2,59	54	2,90	3,16	117
Derivaciones (N25)	Español	6,80	4,24	171	6,50	4,19	54	6,03	4,28	117
	LE	3,95	3,05	171	4,19	3,31	54	3,85	2,93	117
Asociaciones colectivas (N5)	Español	6,15	5,20	171	6,39	5,09	54	6,03	5,27	117
	LE	5,02	4,18	171	5,67	4,23	54	4,73	4,14	117
Palabras nucleares (N1)	Español	176,87	31,56	171	168,8	32,51	54	180,60	30,54	117
	LE	121,39	31,34	171	116,72	31,76	54	123,55	21,04	117

Tabla 5.6. Principales variables lingüísticas

Variables sociodemográficas y reticulares (variables independientes)	Total			Hombres			Mujeres		
	Media	Desv.	N	Media	Desv.	N	Media	Desv.	N
Edad	28,73	6,02	171	30,48	7,32	54	27,92	5,15	117
Años estudio	21,74	3,35	171	21,83	0,45	54	21,70	0,31	117
Otros estudios	0,67	0,80	171	0,74	0,78	54	0,64	0,81	117
Estudios del padre	12,04	7,10	133	3,07	1,66	45	2,89	1,57	88
Estudios de la madre	11,73	7,31	132	2,67	1,52	44	2,7	1,43	88
Ocupación madre	4,66	2,46	171	4,94	2,71	54	4,53	2,34	117
Ocupación padre	5,95	3,53	171	6,13	3,69	54	5,86	3,47	117
Ingresos	3,56	1,01	171	3,67	1,26	54	3,50	1,02	117

Tabla 5.7. Principales variables sociodemográficas y reticulares

Variables relacionadas con el aprendizaje de una lengua extranjera	Total			Hombres			Mujeres		
	Media	Desv.	N	Media	Desv.	N	Media	Desv.	N
Exposición a los medios de comunicación en español	20,32	6,76	171	20,54	7,02	54	20,21	6,67	117

Exposición a los medios de comunicación en lengua extranjera	12,16	8,63	171	12,13	1,13	54	12,18	0,81	117
Contactos externos	5,81	1,51	171	5,63	1,59	54	5,90	1,47	117
Lenguas extranjeras que conoce el informante	2,01	1,87	171	2,02	0,14	54	2,01	0,09	117
Suma de las lenguas extranjeras que conocen los familiares	1,78	1,35	171	1,61	1,35	54	1,86	1,35	117
Suma uso diferentes lenguas	1,07	0,94	171	1,09	0,94	54	1,06	0,95	117
Sensación al impartir clases bilingües	1,58	1,24	171	1,78	1,33	54	1,49	1,19	117
Estudios en el extranjero	1,99	1,06	171	2,06	1,14	54	1,96	1,03	117

Tabla 5.8. Principales variables relacionadas con el aprendizaje de una lengua extranjera

5.4. Análisis bivalente

A través de los análisis bivariantes, se pretende conocer la relación entre el conjunto de datos que ofrecen dos variables. Además de los análisis exploratorios, se han incluido análisis descriptivos e inferenciales como la comparación entre medias mediante pruebas paramétricas y la correlación lineal bivariada. La representación gráfica de la posible relación de tipo lineal entre dos variables se ha realizado mediante diagramas de dispersión. Este tipo de gráficos muestra mediante puntos el valor de una variable en el eje de abscisas y el de la otra en el eje de coordenadas. La línea que dibujan esos puntos marca la tendencia de esas variables.

La recta de regresión sirve para predecir el valor de la variable dependiente (x) a través de la variable independiente (y). Estos datos permiten conocer si la correlación entre dos variables es significativa y, por consiguiente, si es oportuno utilizar esas variables como factores predictores en los análisis de regresión multivariante ulteriores. Por ello, se ha examinado el comportamiento de las variables lingüísticas (índice corregido de descentralización léxica e índice corregido de identificación de tecnicismos) entre sí y se han marcado los resultados según la información aportada por una tercera variable, como por ejemplo, el tipo de centro educativo para predecir la posible relación entre estas tres variables.

Las pruebas paramétricas ayudan a conocer si existen diferencias significativas entre dos o más conjuntos de datos. La *t* de Student se aplica en los casos donde se compara el promedio de dos grupos de datos y establece como hipótesis nula que ambos grupos pertenecen a una misma población si no existen diferencias entre las medias. La prueba ANOVA o análisis de la varianza se calcula para comparar tres o

más conjuntos de datos, analizar sus dispersiones y varianzas e identificar subconjuntos con promedios iguales. La hipótesis nula establece la igualdad de los grupos e indica que las diferencias se deben al azar. La hipótesis alternativa muestra que existen diferencias en la media de al menos uno de los grupos. Antes de calcular estas pruebas estadísticas, se ha establecido como requisito previo que cada uno de los grupos presente una distribución normal y una varianza homogénea.

Una vez que se conoce que existen diferencias entre los grupos poblaciones a través del análisis de la varianza, se aplican dos pruebas *post hoc*, Tukey y Scheffé, con objeto de identificar subconjuntos homogéneos de medias que no se diferencian de forma significativa²³⁷. Por un lado, el método de la diferencia significativa honesta (Honestly Significant Difference) o método de Tukey realiza todas las combinaciones por pares posibles utilizando el rango estudentizado. Por otro lado, el método Scheffé lleva a cabo comparaciones múltiples por pares con la distribución muestral *F*.

5.4.1. Relaciones entre variables lingüísticas

5.4.1.1. Variables lingüísticas en lengua extranjera

En primer lugar, se presentan análisis exploratorios para conocer si existe una relación aparente entre el caudal léxico y la capacidad de identificación de tecnicismos y, a continuación, le siguen análisis descriptivos, correlacionales y comparativos. El diagrama de cajas muestra una estrecha relación entre el índice corregido de descentralización léxica y el índice corregido de identificación de tecnicismos en lengua extranjera, lo cual supone que a mayor número de palabras proporcionadas por los informantes mejor es su conocimiento de tecnicismos. Existe una diferencia considerable entre la media de los dos primeros grupos (0,43 y 0,45) y la del tercer y cuarto grupo (0,58 y 0,61). El último grupo baja el promedio del caudal léxico en 0,56

²³⁷ En la codificación de los datos el investigador puede considerar que cada especialidad del máster da lugar a medias diferenciadas del ICoDD. La prueba de ANOVA indica que las varianzas poblaciones son distintas pero no se conocen qué especialidades presentan medias estadísticamente diferentes. A través de las pruebas Tukey y Scheffé se han establecido subconjuntos homogéneos de especialidades para el ICoDD y el ICIT en español y en lengua extranjera. De esta forma, la posible incidencia de la variable de la especialidad sobre la prueba lingüística se realiza según una clasificación fundamentada en análisis estadísticos en detrimento de la que pudiera realizar el investigador según áreas que considere afines.

debido en parte a que solo existen tres informantes que hayan superado el rango del 75 %.

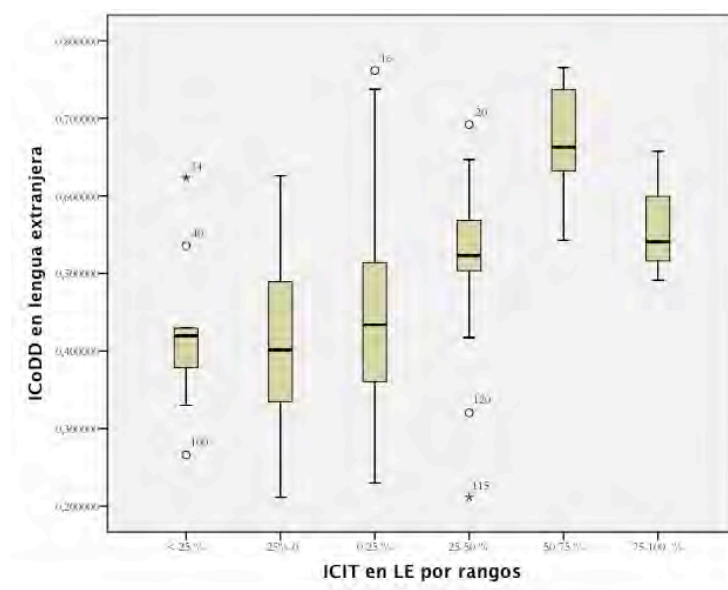


Gráfico 5.10. Diagrama de cajas de la relación entre el ICoDD y el ICIT en LE

Los *outliers* se distribuyen en los cuatro primeros rangos y corresponden a dos tipos de datos diferenciados: informantes cuyo caudal léxico es bastante superior a la capacidad de identificación de tecnicismos para ese rango (casos 16, 20, 34 y 40) y viceversa, es decir, informantes cuyo ICoDD es inferior al que le correspondería según la media del ICIT (100, 115 y 120). Asimismo, es posible diferenciar los *outliers* de los *outliers extremos* porque los primeros están marcados por un círculo (16, 20, 40, 100 y 120) y los segundos por una estrella (34 y 115).

Los casos que superan la media corresponden a las especialidades de Lengua Extranjera Inglés (16 y 20), Lengua Extranjera Francés (34) y Lengua Castellana y Literatura (40). En cambio, los sujetos que destacan por su bajo caudal léxico cursan las especialidades de Orientación Educativa (100) y Hostelería y Turismo (115 y 120). Se presupone que en general los estudiantes de especialidades lingüísticas cuentan con mayor capacidad léxica en lengua extranjera. En cambio, los dos estudiantes que cursan la especialidad de Hostelería y Turismo son diplomados en Turismo y han realizado asignaturas obligatorias durante la carrera, por lo que los resultados se esperaban mejores. *A priori*, esos datos arrojan información sobre la influencia del tipo de especialidad del máster que se analizará en el apartado siguiente.

Entre las hipótesis iniciales, no se ha planteado la posibilidad que el tipo de centro educativo donde el informante ha realizado sus estudios de educación primaria y secundaria incida en el vocabulario activo ni pasivo del informante, por lo que esa variable no se analiza en detalle. No obstante, en la siguiente relación entre el ICIT y el ICoDD en lengua extranjera se ha establecido la diferenciación de los informantes según el tipo de centro donde ha cursado sus estudios de enseñanza secundaria obligatoria y postobligatoria. Se pretende demostrar la importancia de la selección de las variables en cada estudio con independencia de las variables tradicionales, ya que el tipo de centro educativo pierde significancia en estudiantes que han finalizado estudios universitarios.

El gráfico de dispersión muestra una predominancia de los informantes que han estudiado en centros públicos con respecto a los centros concertados y privados respectivamente. La nube de puntos que corresponde al centro público se distribuye de forma homogénea en los dos ejes, es decir, tanto en el caudal léxico como en la capacidad de identificación de tecnicismos. El promedio de respuestas en el ICoDD en este caso es de 0,463 y del ICIT es de 3,021. En cuanto a los resultados del centro concertado, tan solo dos hablantes superan los seis puntos en la prueba del ICIT y otros dos están en la parte inferior con unos resultados mínimos. Las medias de ambos índices son muy similares: 0,479 y 3,061 respectivamente. Por último, los resultados del centro privado se sitúan ligeramente sobre la media en el ICIT con 3,692 pero bajo la media en el ICoDD con un índice de 0,426.

Se observa una relación positiva entre el ICIT y el ICoDD ya que ambos índices aumentan de forma progresiva dibujando, a grandes rasgos, una línea ascendente en el gráfico de dispersión. Se ha comprobado que la correlación es significativa al nivel 0,01 (bilateral) y el coeficiente de correlación de Pearson es de 0,438. De esta forma, se constata que cuanto mayor es el léxico disponible de un informante mayor será su capacidad para identificar el léxico especializado y viceversa.

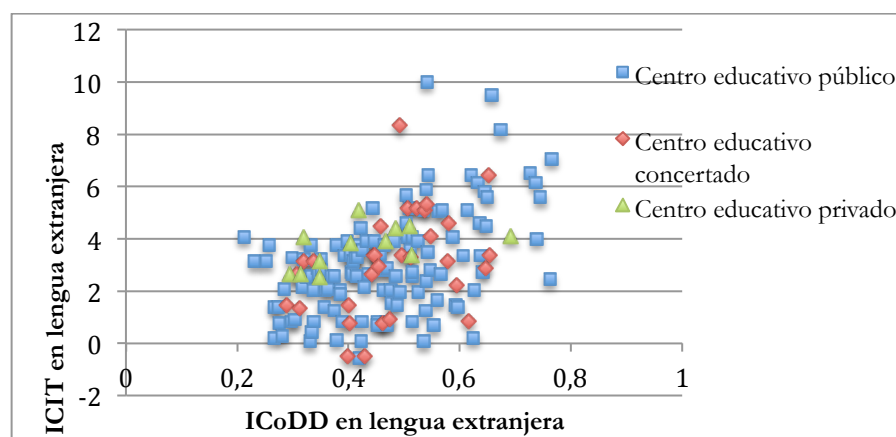


Gráfico 5.11. Diagrama de dispersión del ICoDD y el ICIT en lengua extranjera según el tipo de centro educativo

5.4.1.2. Variables lingüísticas en lengua española

Los datos exploratorios muestran que la media del caudal léxico en los cinco rangos mostrados para el índice corregido de identificación de tecnicismos es muy similar: 0,39 en el primero; 0,47 en el tercero (ya que no existe ningún informante en el segundo tramo); 0,48 en el cuarto y 0,51 en el quinto. Aunque el léxico especializado no influye de forma tan notoria en el caudal léxico en español como lo hace en lengua extranjera, la relación muestra una progresión constante de forma que conforme aumenta el léxico especializado lo hace el ICoDD en esa lengua. Esa similitud de los datos viene a reforzar la hipótesis que en lengua materna el factor con más peso es el nivel de instrucción razón por la cual las diferencias entre los informantes son menores.

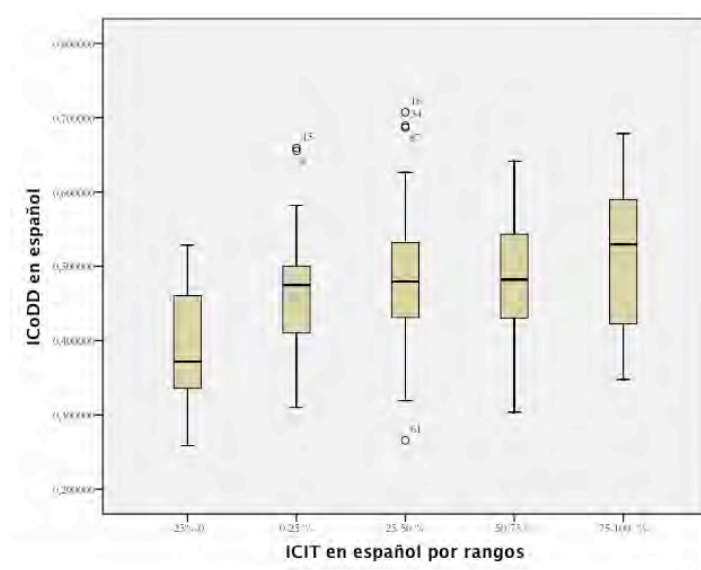


Gráfico 5.12. Diagrama de cajas que del ICoDD en español en relación con el ICIT por rangos

En esta ocasión, no existe ningún caso de *outlier* extremo sino de valores atípicos moderados. Al igual que en lengua extranjera, los informantes 16 y 34 cuentan con una media del ICoDD superior al rango del ICIT al que pertenecen por lo que se constata que son informantes muy productivos tanto en español como en lengua extranjera. Existen otros tres sujetos con un caudal léxico superior al rango del ICIT en el que se sitúan y que están matriculados en las especialidades de Lengua Extranjera Inglés (15), Lengua Extranjera Francés (34) y Música (87). Se podría decir que las especialidades que sobresalen según los resultados del léxico especializado son las dos de lenguas extranjeras, lo que lleva a pensar que el conocimiento de más de un idioma refuerza no solo el vocabulario especializado en la lengua extranjera sino también en la lengua materna. En el extremo opuesto se sitúa el informante 61 que cursa la especialidad de Geografía e Historia.

En lengua española, los datos difieren ligeramente ya que los resultados del ICoDD y del ICIT son superiores con independencia del tipo de centro educativo. Los informantes de los centros educativos públicos se concentran en torno a la media pero también cuentan con algunos de los índices más altos, sobre todo en el ICoDD. La media de este índice es de 0,479 y del ICIT es de 5,477. Los centros concertados se desplazan levemente a la derecha contando con una media algo superior del ICoDD, concretamente 0,481, y del ICIT, 5,873. En cuanto a los resultados de los centros privados, se localizan en la parte central superior y presentan una media del ICoDD algo inferior (0,452) y el mayor promedio del ICIT de los tres tipos de centros (6,516).

En síntesis, los datos del léxico en español no se ven influidos de forma significativa por el tipo de centro educativo donde los informantes han cursado los estudios de enseñanza secundaria. Los datos del gráfico se han contrastado con pruebas de correlaciones bivariadas y de análisis de la varianza y no muestran resultados significativos. Esa similitud de los datos se repite tanto en español como en lengua extranjera con lo que esta variable no se considera relevante para explicar el comportamiento lingüístico de los hablantes. En cambio, se muestra una relación lineal entre las dos variables lingüísticas.

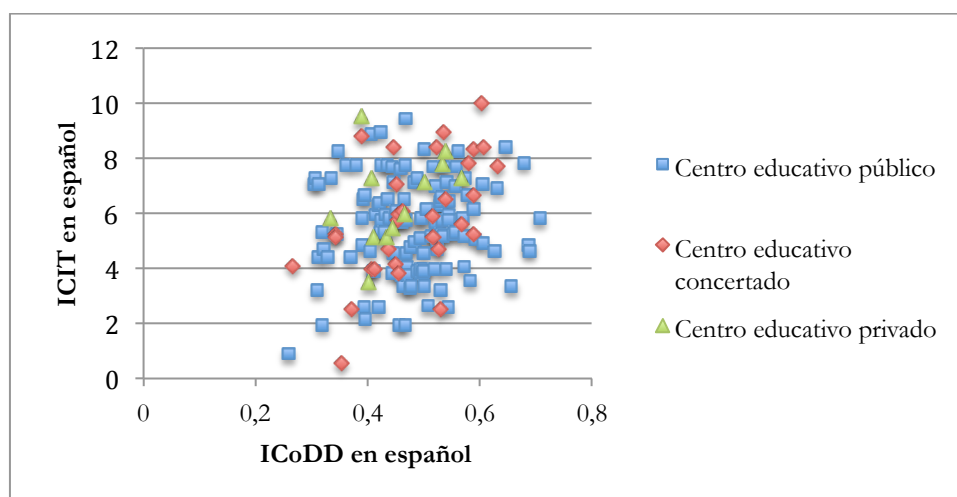


Gráfico 5.13. Diagrama de dispersión del ICoDD y el ICIT en español según el tipo de centro educativo

5.4.1.3. Caudal léxico en lengua extranjera y en lengua española

Una vez que se ha comprobado la relación positiva entre las variables lingüísticas en una misma lengua, se pasa a verificar si existe tal correlación entre cada variable lingüística en lengua materna y lengua extranjera. En este apartado se analiza la relación entre el caudal léxico en lengua española y en lengua extranjera. El objetivo es comprobar la tendencia lineal de una variable con respecto a la otra²³⁸ con independencia que el ICoDD se vea alterado por el número de palabras e informantes.

El diagrama de dispersión muestra una relación positiva entre el caudal léxico en lengua extranjera y en lengua española de forma que al aumentar un índice en un idioma también lo hace en el otro aunque no sea en proporción constante. La correlación es significativa a un nivel de 0,01 y el coeficiente de Pearson toma un valor de 0,633 lo que muestra la fuerte relación entre ambas variables. En cuanto al idioma de la prueba, parece no tener relación con ninguna de las variables lingüísticas ya que tanto la lengua inglesa como la francesa aparecen en los niveles superiores,

²³⁸ Cabe precisar que la relación entre el índice corregido de descentralización léxica, el índice de amplitud léxica y el índice corregido de amplitud léxica mantienen una relación casi directa de forma que cuando una de ellas aumenta la otra lo hace de forma casi constante. Por tanto, al igual que se pueden comparar las medias de palabras en una lengua se puede comprobar el ICoDD.

centrales e inferiores. Asimismo, se ha calculado la correlación bilateral entre el idioma y las variables lingüísticas y no ha sido significativa²³⁹.

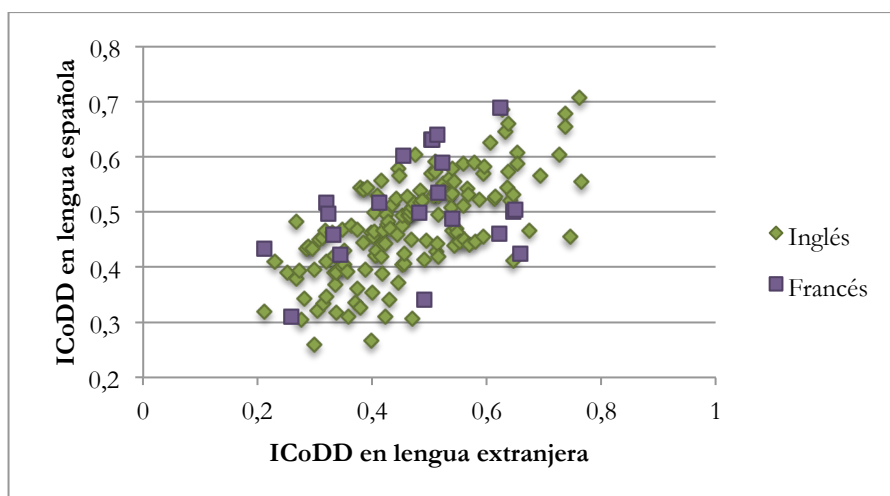


Gráfico 5.14. Diagrama de dispersión del ICIT en español y en lengua extranjera según el idioma de la prueba

Esta relación viene a reforzar la idea de que cuanto mayor es la capacidad léxica en un idioma también lo es en otra lengua. Los informantes con un mayor caudal léxico en español han repetido ese modelo en lengua extranjera. No obstante, existen diferencias entre los resultados porque en lengua extranjera intervienen factores relacionados con el aprendizaje de una lengua extranjera cuya influencia en lengua española es más leve.

5.4.1.4. Léxico especializado en lengua extranjera y en lengua española

La hipótesis planteada en relación con el léxico pasivo pasa por considerar que el factor con mayor influencia en lengua española es el grado de instrucción de los informantes. Se trata de un léxico que el informante es capaz de identificar y asociar pero que no necesariamente tiene que estar disponible ni formar parte del vocabulario activo del hablante. Si un informante escucha la palabra *ribosoma* seguramente la asociará al centro de interés del cuerpo humano por similitud a otras palabras como *cromosoma*. En cambio, si no es especialista en la materia difícilmente

²³⁹ Se ha comprobado que la correlación entre la lengua de la prueba y el léxico especializado es significativa a nivel 0,01 con un coeficiente de Pearson de 0,294. Asimismo, en los resultados de la prueba de *t* de Student se asume la hipótesis alternativa. El *p* valor es menor de 0,05 lo que quiere decir que existen diferencias entre las varianzas. De ahí que esta apreciación sea tenida en cuenta en los análisis multivariantes.

sabría utilizarlo en la situación adecuada y tampoco será una de las primeras palabras que le venga a la mente en una prueba de disponibilidad léxica.

En lengua extranjera, se ha mostrado que la relación del léxico especializado y el caudal léxico es más estrecha ya que los informantes con mayor capacidad léxica son también los que obtienen mejores resultados en la prueba del léxico especializado. Ese léxico especializado está más influido por factores relacionados con el aprendizaje de una lengua extranjera y, en menor medida en lengua española.

La correlación resultante entre el léxico especializado en español y en lengua extranjera es significativa a un nivel de 0,01, como en el caso anterior, pero el coeficiente de correlación de Pearson es más débil, concretamente 0,296. De esta forma, se constata que los resultados del índice corregido de descentralización léxica y del índice corregido de identificación de tecnicismos se trasladan de una lengua a otra aunque guardan diferentes grados de relación.

En cuanto a la variable que marca los resultados, muestra la movilidad del informante para trasladarse al extranjero en busca de trabajo²⁴⁰. Los datos apuntan a que en general los estudiantes muestran esa movilidad porque en España la mayoría de las personas que emigran están altamente cualificadas²⁴¹. No obstante, se parte también del supuesto que la razón de aquellos estudiantes que se muestran más reacios a buscar trabajo en el extranjero puede deberse a que no tienen un buen dominio de la lengua extranjera. El gráfico refleja la tendencia de la posible emigración de los informantes (155 frente a 16 sujetos) y sitúa a aquellos que no desean trasladarse al extranjero en la franja intermedia y baja que indica unos resultados débiles del léxico especializado. Concretamente se ha realizado la correlación del ICIT con esta variable y es significativa en el caso de la lengua extranjera a un nivel de 0,05 y con un coeficiente de correlación de 0,168.

²⁴⁰ Corresponde a la pregunta 18, apartado tercero, del cuestionario del Anexo IV.

²⁴¹ El estudio de Alaminos Chica y Santacreu Fernández (2010) muestra que a partir de la entrada de España en la Comunidad Europea se produce un incremento lineal de las migraciones de trabajadores cualificados que reemplaza la de españoles no cualificados no cualificados. Uno de los factores que destaca en este fenómeno es el interés por mejorar las expectativas futuras de trabajo de los trabajadores cualificados y su mejor conocimiento de idiomas extranjeros. Disponible en: http://rua.ua.es/dspace/bitstream/10045/23304/1/2010_Alaminos_Santacreu_Papers.pdf [Fecha de última consulta el 30 de marzo de 2015].

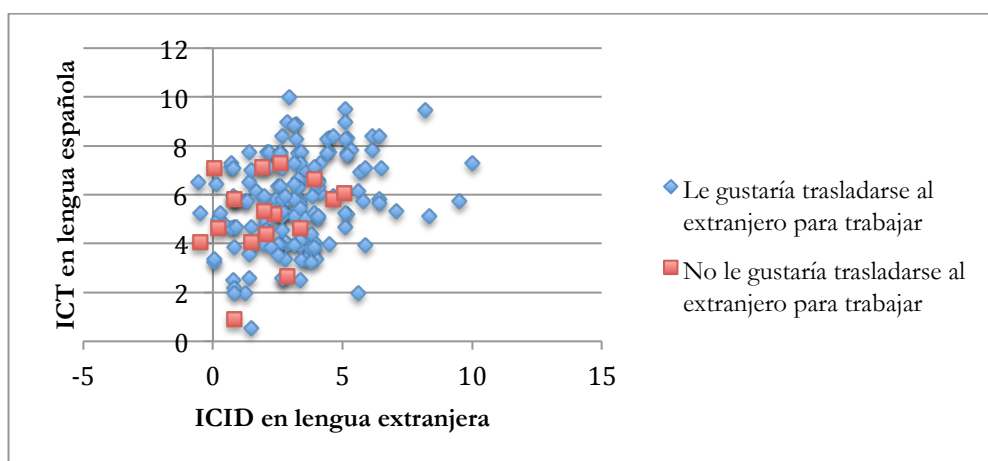


Gráfico 5.15. Diagrama de dispersión del ICIT en lengua extranjera y en español según la movilidad del informante

5.4.2. Efecto de las variables socioculturales

Se parte de la hipótesis que la clase social del estudiante incide de forma sutil en el caudal léxico y en el léxico especializado del informante ya que otros factores, como el tipo de estudios previos, cobran mayor importancia. Esa influencia será menor en lengua española donde el grado de instrucción hace que se homogenicen los resultados. El estatus social puede tener mayor relevancia en lengua extranjera ya que los ingresos familiares así como el tipo de estudios y ocupación de los padres pueden influir en el aprendizaje de ese idioma por diversos motivos: realización de estancias y viajes en el extranjero, asistencia a academias especializadas, entre otros aspectos.

En este apartado, se han comparado las medias de las variables que definen mejor la clase social de un individuo: la ocupación y estudios de los padres y los ingresos familiares. Tras realizar la comparación de medias a través del cálculo de Anova de un factor y el cotejo de las varianzas, se constata que el promedio de los grupos de cada variable es ligeramente diferente aunque en lengua española los grupos proceden de poblaciones con medias iguales. En cambio, la correlación de Pearson, realizada entre variables lingüísticas y variables socioculturales, es significativa al nivel 0,05 (bilateral). A modo de ejemplo, se presentan los resultados de la comparación de medias del léxico especializado en lengua extranjera y del caudal léxico en lengua materna con respecto a los estudios del padre.

5.4.2.1. Estudios del padre y léxico especializado en lengua extranjera

El índice corregido de identificación de tecnicismos en lengua extranjera mantiene una correlación significativa a nivel 0,05 con los estudios alcanzados por el padre y presenta un coeficiente de correlación de Pearson de 0,170. Esa relación positiva se observa en el promedio de respuestas del léxico especializado que aumenta en consonancia con los estudios del padre salvo en dos ocasiones: el promedio de los padres que no han realizado ningún tipo de estudios es de 3,005 y de los que tienen la primaria incompleta es de 2,298; del mismo modo la media del ICIT de los padres con Bachillerato o un grado medio de FP es de 3,396 y de los que han realizado una diplomatura o un grado superior de FP es de 3,150.

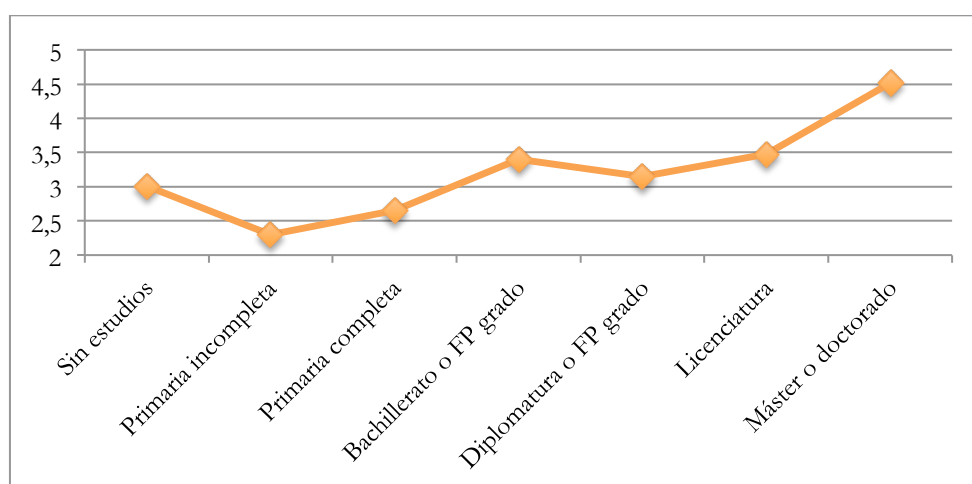


Gráfico 5.16. Distribución de la media del índice corregido de identificación de tecnicismos con respecto al grado de estudios del padre

Los datos descriptivos muestran que el máximo de todas las respuestas lo ha realizado un informante cuyo padre no ha realizado ningún tipo de estudios. Las variables relacionadas con el aprendizaje de una lengua extranjera tienen una gran relevancia en ese estudiante ya que cursa la especialidad del máster en lengua extranjera, ha realizado una estancia en el extranjero y conoce tres idiomas extranjeros. El dato mínimo corresponde a un informante de la especialidad de Matemáticas que habla un solo idioma con un nivel B1. El hecho de que el padre haya finalizado los estudios de Primaria no es suficiente para determinar un mayor conocimiento pasivo del léxico especializado ya que la incidencia en este caso de las variables relacionadas con el aprendizaje de una lengua es mínima.

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95 %		Mínimo	Máximo
					Límite inferior	Límite superior		
Sin estudios	16	3,005	2,531	0,633	1,656	4,354	-0,49	10,00
Primaria incompleta	12	2,298	1,152	0,333	1,566	3,030	0,84	5,11
Primaria completa	43	2,650	1,839	0,280	2,084	3,216	-0,56	8,33
Bachillerato o FP grado medio	33	3,396	2,028	0,353	2,677	4,115	-0,49	9,52
Diplomatura o FP grado superior	30	3,150	1,991	0,363	2,407	3,893	0,21	7,07
Licenciatura	33	3,474	1,103	0,195	3,076	3,871	1,40	6,44
Máster o doctorado	4	4,515	2,647	1,323	0,303	8,727	2,03	8,19

Tabla 5.9. Datos descriptivos de los estudios del padre con respecto al índice corregido de identificación de tecnicismos en lengua extranjera

En cuanto a los resultados de ANOVA, el valor del estadístico de prueba ($F = 0,737$) no es claramente distinto de 1 y presenta una significación de 0,620. Para completar los análisis, se han realizado las pruebas *post hoc* de Tukey y Scheffé y han mostrado la homogeneidad de las diferentes variantes de este factor. Por lo tanto, se confirma la hipótesis nula de igualdad de medias que afirma que los diferentes grupos pertenecen a una misma población y que las diferencias se deben al azar. Asimismo, se concluye que existe una correlación positiva entre la variable de la formación del padre y del léxico especializado sin que los grupos muestren diferencias significativas.

5.4.2.2. Estudios del padre y caudal léxico en lengua española

Al igual que en el apartado anterior, la relación entre estas dos variables es positiva de forma que conforme aumenta el nivel de estudios alcanzado por el padre mayor es la capacidad léxica del informante. La correlación bivariada es significativa a un nivel de 0,05 y el resultado del coeficiente de Pearson es de 0,166. En cambio, el promedio del índice corregido de descentralización en lengua española es similar en todas las categorías de la variable ya que el grupo con menor media es de 0,447, correspondiente a los padres que no han terminado los estudios primarios, y con el grupo mayor promedio lo presentan los padres que han realizado máster o doctorado con un índice de 0,512. Existen tres grupos cuya media baja ligeramente con respecto al grupo anterior y, por tanto, con menor formación académica: es el caso de la

primaria incompleta, los estudios de Bachillerato o FP de grado medio y la licenciatura.

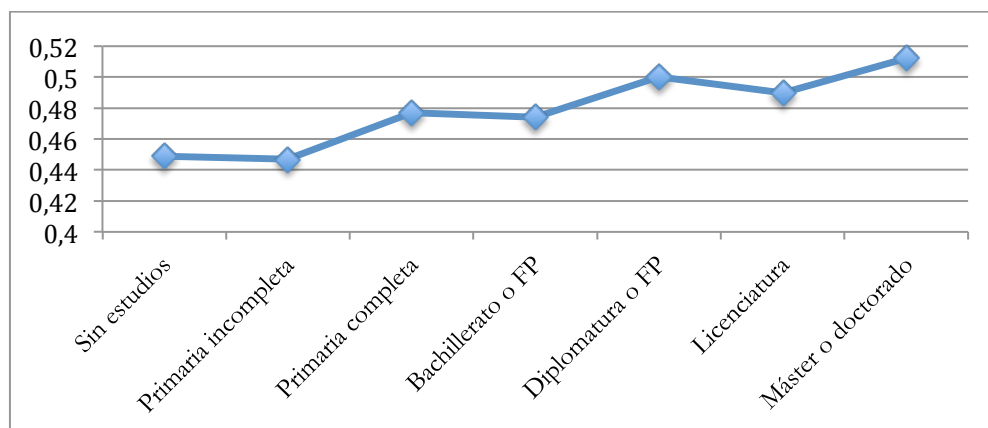


Gráfico 5.17. Distribución de la media del índice corregido de descentralización léxica en español con respecto al grado de estudios del padre

En lengua española el valor máximo (0,707) y mínimo (0,259) lo presentan informantes cuyos padres han completado solamente los estudios primarios lo cual adelanta que el mayor o menor ICoDD en cada grupo puede deberse al azar. Por orden de rango en los máximos valores encontrados, le seguiría el segundo grupo que hace referencia a padres con una diplomatura o FP de grado medio, luego se sitúan los que han realizado una licenciatura, Bachillerato o FP de grado medio, máster o doctorado, sin estudios y con primaria incompleta. Por tanto, los máximos registrados entre los informantes cuyos padres no tienen estudios (0,590) es muy similar al de padres que han realizado el máster o doctorado (0,605).

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
					Límite inferior	Límite superior		
Sin estudios	16	0,449	0,082	0,021	0,405	0,493	0,266	0,590
Primaria incompleta	12	0,447	0,081	0,023	0,396	0,499	0,319	0,541
Primaria completa	43	0,477	0,097	0,015	0,447	0,507	0,259	0,707
Bachillerato o FP grado medio	33	0,474	0,073	,013	0,448	0,500	0,310	0,641
Diplomatura o FP grado superior	30	0,500	0,100	0,018	0,462	0,537	0,306	0,689
Licenciatura	33	0,490	0,093	0,016	0,456	0,524	0,327	,0659
Máster o doctorado	4	0,512	0,073	0,036	0,396	0,628	0,443	0,605
Total	171	0,479	0,090	0,007	0,465	0,493	0,259	0,707
Modelo	Efectos fijos		0,090	0,007	0,465	0,493		

Efectos aleatorios	0,007	0,462	0,496
-----------------------	-------	-------	-------

Tabla 5.10. Datos descriptivos de los estudios del padre con respecto al índice corregido de identificación de descentralización léxica en lengua extranjera

En cuanto a los valores mínimos de cada grupo, existe una cierta progresión con respecto al nivel de estudios de los padres de forma que el mínimo más alto corresponde a los informantes cuyos progenitores cuentan con estudios de máster o doctorado (0,443) y licenciatura (0,327). En cambio, el siguiente en la lista es el grupo de padres con la primaria incompleta (0,319) lo que vuelve a mostrar que los mínimos más altos o más bajos se pueden encontrar prácticamente en cualquier grupo.

El ANOVA es prácticamente igual a uno ($F = 1,003$) con un nivel de significación de 0,425. La similitud en la varianza indica que no tiene sentido buscar subgrupos mediante pruebas *post hoc* confirmando la hipótesis nula de que las poblaciones cuentan con medias iguales. Asimismo, la correlación de Pearson es muy similar en lengua materna y en lengua extranjera lo que refuerza la idea previa que las variables relacionadas con la clase social tiene una incidencia débil. Se pronostica que la influencia de las variables relacionadas con el aprendizaje de una lengua tenga un impacto mucho más considerable en las variables lingüísticas en el idioma extranjero y, en menor medida, en lengua materna.

5.4.3. Efecto de las variables relacionadas con el dominio de una lengua

Según una de las hipótesis iniciales, las variables presentan una influencia mayor en los resultados del léxico activo que en los resultados de léxico pasivo. Las variables que se presentan en este apartado son la especialidad del máster, el conocimiento de otros idiomas extranjeros, el uso de la lengua extranjera en diferentes contextos, los hábitos de lectura, la exposición a los medios de comunicación y la motivación del futuro profesorado para impartir clases en enseñanza bilingüe. Además, cabe destacar que la mayoría de los factores están más

relacionados con el aprendizaje de una lengua extranjera por lo que su incidencia en español²⁴² es menor.

5.4.3.1. Especialidad del máster

Una vez expuesta la relación entre especialidad y titulación en los análisis univariantes, se muestran los datos exploratorios por especialidades y áreas con respecto al índice corregido de descentralización léxica en lengua extranjera. El número de informantes por áreas se distribuye de la siguiente forma: 74 en el área social-lingüística, 50 en el área científico-tecnológica, 20 en el área artística y 27 en el área de formación profesional. Se presupone que el tipo de especialidad influye de forma considerable en el léxico en lengua extranjera ya que marcan los estudios previos de los informantes y, por ende, las asignaturas en lengua extranjera realizadas durante los estudios universitarios. En cambio, esta hipótesis no se mantiene en el caso de la lengua materna donde la instrucción universitaria viene a reforzar el léxico con unidades léxicas más especializadas que los hablantes desconocían antes de acceder a la diplomatura, ingeniería o licenciatura de que se trate.

Índice corregido de descentralización en lengua extranjera por áreas		Estadístico	Error típ.
Social-lingüística	Media	0,54773182	0,011364958
	Mediana	0,539427	
	Varianza	0,01	
	Desv. típ.	0,097765062	
	Rango	0,43528	
	Amplitud intercuartil	0,145547	
	Asimetría	0,295	0,279
	Curtosis	-0,246	0,552
Científico-tecnológica	Media	0,38265252	0,013424232
	Mediana	0,371646	
	Varianza	0,009	
	Desv. típ.	0,094923657	
	Rango	0,441458	
	Amplitud intercuartil	0,13988	
	Asimetría	0,58	0,337
	Curtosis	0,104	0,662
Artística	Media	0,47340495	0,02280769

²⁴² Vid. apartado 5.3 para una revisión concreta del tipo de variable.

	Mediana	0,4745965	
	Varianza	0,01	
	Desv. típ.	0,10199909	
	Mínimo	0,26762	
	Máximo	0,637094	
	Rango	0,369474	
	Amplitud intercuartil	0,149316	
	Asimetría	-0,237	0,512
	Curtosis	-0,468	0,992
Formación Profesional	Media	0,39485237	0,017917504
	Mediana	0,399289	
	Varianza	0,009	
	Desv. típ.	0,093102081	
	Rango	0,420746	
	Amplitud intercuartil	0,109865	
	Asimetría	0,603	0,448
	Curtosis	0,982	0,872

Tabla 5.11. Datos exploratorios del área del máster con respecto al ICoDD en lengua extranjera

Se aprecian notables diferencias en la media por áreas, pasando de 0,548 en el área social-lingüística hasta un 0,383 en el área científico-tecnológica. Los datos más compactos se encuentran en formación profesional con un IQR de 0,110 seguidos del área científico-tecnológica con un IQR de 0,140 y posteriormente del área sociolingüística y artística con un IQR de 0,146 y 0,149 respectivamente.

Con respecto al rango, se constata que las cuatro áreas están comprendidas en la media presentando una distribución normal o mesocúrtica a excepción del área de formación profesional:

- El área social-lingüística presenta una curtosis de -0,246 y un rango de 0,435.
- El área científico-tecnológica muestra una curtosis ligeramente positiva de 0,106 y un rango de 0,442.
- El área artística cuenta también con una curtosis negativa de -0,468 y un rango de 0,370.
- El área de formación profesional presenta una curtosis platicúrtica bastante más elevada, con 0,982 y un rango de 0,42.

El siguiente diagrama de cajas muestra la distribución explicada anteriormente. Asimismo, se destacan tres *outliers* que corresponden a tres informantes cuyos datos superan la media de su conjunto que se corresponden con las siguientes áreas:

- En el área científico-tecnológica sobresale el informante 146 con un promedio del ICoDD en lengua extranjera de 0,6533. Cabe destacar que este área cuenta con la media global más baja con un índice de 0,383 ya que engloba carreras más técnicas donde como norma general no se imparte ninguna asignatura obligatoria en lengua extranjera. En este caso, el sujeto ha cursado estudios de Arquitectura y actualmente estudia la especialidad del máster denominada Tecnología, Informática y Procesos Industriales. En su caudal léxico han tenido influencia otros factores relacionados con el aprendizaje de lengua extranjera como los estudios fuera de España.
- En el área de Formación y Orientación Laboral destacan los informantes 109 y 111 cuyas medias de ICoDD son de 0,632 y 0,594 respectivamente mientras que la media del grupo es de 0,395. Ambos son Licenciados en Derecho y cursan la especialidad de Formación y Orientación Laboral y han realizado estancias en el extranjero.

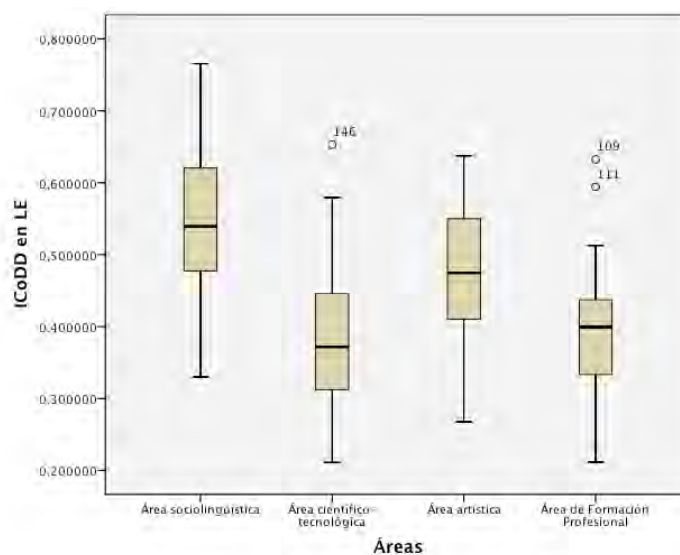


Gráfico 5.18. Diagrama de cajas de la distribución por áreas según el ICoDD en lengua extranjera

A continuación, se muestra la media concreta de cada especialidad por orden decreciente del ICoDD. Se confirma que las especialidades de lenguas extranjeras ocupan los primeros rangos y la especialidad con menos caudal léxico corresponde al

grupo del área científico-tecnológica, Matemáticas. La especialidad Lengua Extranjera Inglés se desmarca ligeramente del Francés y de Lengua y Literatura por lo que quizás su incidencia sea mayor en los análisis de regresión del apartado 5.5.2.1.

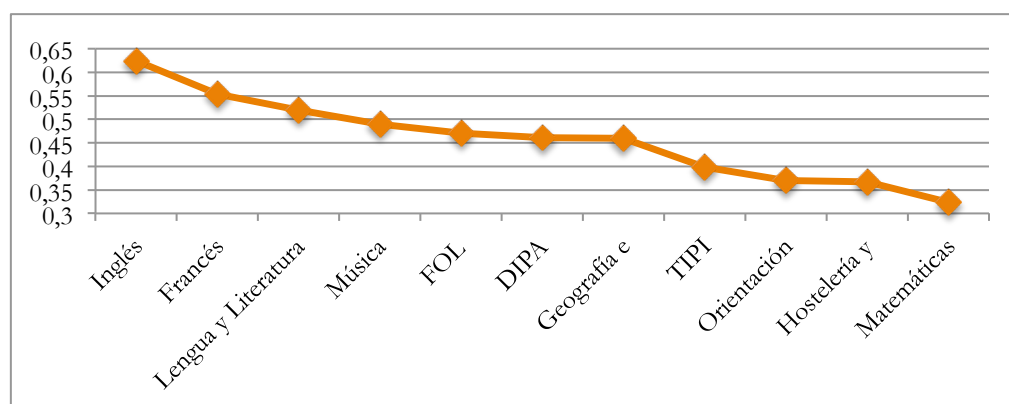


Gráfico 5.19. Promedio del ICoDD en lengua extranjera según la especialidad del máster

Cabe destacar que Hostelería y Turismo ocupa el penúltimo lugar y, sin embargo, nueve de los diez estudiantes han realizado la diplomatura en Turismo y uno la licenciatura en Publicidad y Relaciones Públicas. Cabría plantearse cuál es el nivel que han adquirido los estudiantes de Turismo durante la carrera ya que han cursado asignaturas obligatorias en lengua extranjera. Según el plan de estudios, podría situarse tras la especialidad de Lengua y Literatura ya que los estudiantes que forman parte de esta especialidad han realizado una licenciatura en lugar de una diplomatura y también han cursado asignaturas en lengua extranjera. A partir de ahí, el resto de carreras universitarias de las distintas especialidades no cuentan con materias de carácter obligatorio en lengua extranjera.

Especialidad	Media del ICoDD en lengua extranjera
Inglés	0,625
Francés	0,554
Lengua y Literatura. Latín y Griego	0,520
Música	0,489
Formación y Orientación Laboral (FOL)	0,470
Dibujo, Imagen y Artes Plásticas (DIAP)	0,461
Geografía e Historia	0,460
Tecnología, Informática y Procesos Industriales (TIPI)	0,398
Orientación	0,370
Hostelería y Turismo	0,367
Matemáticas	0,323

Tabla 5.12. Análisis exploratorio del ICoDD en lengua extranjera por especialidades

A continuación, se presenta el análisis de la varianza de la especialidad del máster que ayuda a conocer si se trata de una variable independiente significativa. Para ello, se han comparado las medias de las especialidades con respecto a las cuatro posibles variables dependientes: el índice corregido de descentralización léxica en español y en lengua extranjera y el índice corregido de identificación de tecnicismos en ambas lenguas.

El estadístico F es diferente a uno para cualquier nivel de significación en los cuatro análisis. Por tanto, se rechaza la hipótesis nula de la igualdad de medias que en la muestra analizada quiere decir que existen diferencias entre las distintas variantes o especialidades del máster. Con respecto a los idiomas del estudio, se observa que el estadístico F es considerablemente superior en el caso del caudal léxico en lengua extranjera y en lengua español lo cual se explica la desigualdad en el léxico aportado por informantes de unas especialidades con respecto a otras. No obstante, el valor de F en el léxico especializado es muy similar en español y en lengua extranjera. Ello conlleva que las diferencias entre las especialidades cuando los informantes han hecho la prueba de identificación léxica sean menos dispares.

			Suma de cuadrados	gl	Media cuadrática	F	Sig.
ICoDD lengua extranjera	en	Inter-grupos	1,375	10	0,138	18,863	0,000
		Intra-grupos	1,166	160	0,007		
		Total	2,542	170			
ICoDD español	en	Inter-grupos	0,501	10	0,05	9,23	0,000
		Intra-grupos	0,868	160	0,005		
		Total	1,369	170			
ICIT en lengua extranjera		Inter-grupos	127,322	10	12,732	4,353	0,000
		Intra-grupos	468	160	2,925		
		Total	595,323	170			
ICIT español	en	Inter-grupos	131,896	10	13,19	4,929	0,000
		Intra-grupos	428,157	160	2,676		
		Total	560,053	170			

Tabla 5.13. Análisis de la varianza de las especialidades del máster

La clasificación presentada en los análisis anteriores se basa en las especialidades y áreas en las que se dividen las asignaturas según el Decreto 327/2010. En cambio, es posible que los subconjuntos poblaciones muestren otra organización diferente según la prueba lingüística con la que se compare. Este

aspecto se considera más detenidamente en los análisis multifactoriales ya que permiten conocer qué especialidades inciden positiva o negativamente en el vocabulario virtual y en el especializado de los estudiantes.

5.4.3.2. Conocimiento de idiomas

Tomando como referencia los estudios que avalan la flexibilidad cognitiva de los hablantes bilingües, se ha relacionado el número de idiomas extranjeros que conoce un informante con su caudal léxico y con la capacidad de identificación de tecnicismos en lengua extranjera. A continuación, se han introducido datos más concretos sobre el nivel de dominio de los diferentes idiomas para comprobar su efecto sobre el léxico. Los análisis exploratorios muestran que los informantes que hablan una sola lengua extranjera tienen una media de ICoDD en lengua extranjera de 0,419, aquellos que hablan dos lenguas 0,441, tres lenguas 0,535 y cuatro o más lenguas 0,595. Por tanto, existe una relación positiva entre el número de idiomas que conoce un informante y su capacidad léxica en lengua extranjera.

En los tres primeros casos, se puede considerar que la distribución de los datos es normal ya que los coeficientes de asimetría y curtosis oscilan entre $\pm 0,5$. Sin embargo, en el caso de los informantes que conocen cuatro o más lenguas la distribución de la curtosis es leptocúrtica con un coeficiente de asimetría de -0,313 y de curtosis de -0,796. El índice de amplitud intercuartil es bastante bajo como se refleja en las formas alargadas del diagrama de barras y, por consiguiente, muestra una gran concentración de los valores en torno a la media.

Dominio de lenguas extranjeras		Estadístico	Error típ.
Una lengua extranjera	Media	0,41921941	
	Mediana	0,4195265	
	Desv. típ.	0,092254163	
	Rango	0,384729	
	Amplitud intercuartil	0,131482	
	Asimetría	0,303	0,299
	Curtosis	-0,506	0,59
Dos lenguas extranjeras	Media	0,44176355	0,017109519
	Varianza	0,016	
	Desv. típ.	0,126887589	
	Rango	0,554014	

Tres lenguas extranjeras	Amplitud intercuartil	0,191637	
	Asimetría	0,176	0,322
	Curtosis	-0,442	0,634
	Media	0,53522837	0,016195937
	Varianza	0,01	
	Desv. típ.	0,099838459	
	Rango	0,41365	
	Amplitud intercuartil	0,105605	
	Asimetría	-0,045	0,383
	Curtosis	-0,15	0,75
Cuatro lenguas extranjeras	Media	0,59486993	0,031476962
	Varianza	0,014	
	Desv. típ.	0,117776007	
	Rango	0,360307	
	Amplitud intercuartil	0,155433	
	Asimetría	-0,313	
	Curtosis	-0,796	1,154

Tabla 5.14. Análisis exploratorios del conocimiento de lenguas extranjeras con respecto al ICoDD en lengua extranjera

El diagrama de cajas muestra claramente la relación entre el número de idiomas y el caudal léxico de forma que cuantas más lenguas conoce un informante, mayor es su ICoDD en lengua extranjera. Los casos extremos se localizan en la variante de los informantes que conocen tres idiomas. El informante número 8 conoce tres idiomas con un nivel avanzado y cursa la especialidad de Lengua Extranjera Inglés. En el lado opuesto, se sitúan los informantes 101 y 167 ya que destacan por su bajo nivel de ICoDD. Se trata de hablantes que conocen tres lenguas extranjeras con un nivel básico de dominio.

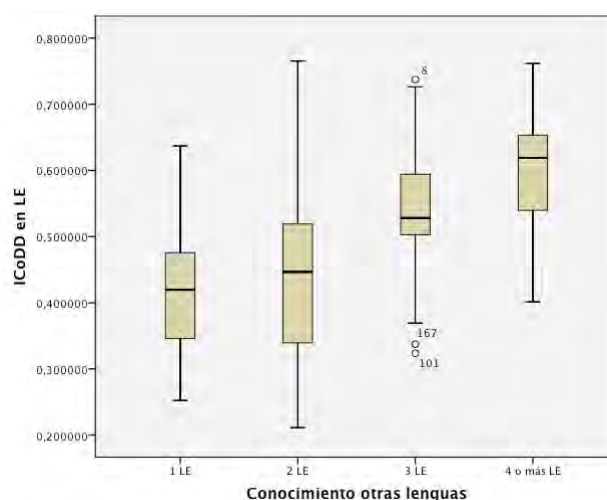


Gráfico 5.20. Diagrama de cajas del conocimiento de otras lenguas sobre el ICoDD en lengua extranjera

En el cuestionario, las preguntas sobre el conocimiento de otras lenguas extranjeras son más específicas y hacen referencia al número de lenguas y nivel de conocimiento alcanzado. Se ha establecido una clasificación en base a los resultados y se han creado tantas categorías como posibilidades han presentado los informantes. Los niveles A1 y A2 se han considerado niveles iniciales, el B1 nivel medio y los niveles B2 y C1 avanzado. No hay ningún informante que haya marcado la casilla del nivel C2 y muy pocos han marcado la casilla del C1.

De esta forma, es posible conocer que de los 64 informantes que hablan una sola lengua, 50 afirman tener un conocimiento de un nivel B1 que es el mínimo establecido en el máster. En el extremo opuesto, se sitúan 14 informantes que conocen cuatro o más lenguas extranjeras que provienen de especialidades en lengua extranjera.

Dominio de lenguas extranjeras	Número de informantes	ICoDD en lengua extranjera ²⁴³	Nivel	Número de informantes	ICoDD en lengua extranjera
Una lengua extranjera	64	0,419	Medio	50	0,395
			Avanzado	14	0,505
Dos lenguas extranjeras	55	0,442	Inicial y medio	22	0,402
			Medio	16	0,424
			Medio y avanzado	17	0,511
Tres lenguas extranjeras	38	0,535	Inicial	13	0,497
			Inicial y medio	10	0,551
			Medio y avanzado	15	0,558
Cuatro o más lenguas extranjeras	14	0,595	Inicial y medio	8	0,591
			Medio y avanzado	6	0,600
Total	171			171	

Tabla 5.15. Distribución de frecuencias según el número de lenguas extranjeras que conoce el informante

Tal y como se puede apreciar en la Tabla 5.14, la diferencia de medias entre los hablantes que conocen una lengua extranjera con un nivel medio y aquellos con un nivel avanzado es notoria, en el primer caso es de 0,395 y en el segundo 0,506. No obstante, cabe destacar que el ICoDD de los hablantes que conocen dos idiomas con niveles inicial o medio desciende a 0,40. El factor del dominio de una lengua influye de forma tan considerable por tratarse de la misma lengua del cuestionario. Se

²⁴³ Se establece la media del índice corregido de descentralización léxica (ICoDD) para los grupos formados por una, dos, tres y cuatro lenguas.

demuestra que si el hablante cuenta con un nivel avanzado su índice de ICoDD será superior.

El siguiente intervalo corresponde al dominio de dos idiomas con nivel medio cuyo ICoDD aumenta hasta 0,42 y le sigue el dominio medio y avanzado de dos idiomas con un 0,51. En cambio, si se trata de un informante que conoce tres idiomas con un nivel básico y medio el ICoDD vuelve a descender ligeramente, concretamente a 0,50 para aumentar de nuevo a la par del dominio de idiomas: 0,55 en el caso de tres idiomas con nivel medio y 0,56 en el supuesto de tres idiomas con nivel medio y avanzado. Por último, los niveles superiores del ICoDD corresponden a los informantes que conocen cuatro o más idiomas, tanto para el nivel inicial y medio (0,59) como para el avanzado (0,60).

En líneas generales, se destaca que cuanto mayor es el dominio de idiomas de un hablante mejor es su ICoDD. Además, el ICoDD aumenta según el nivel adquirido en los idiomas extranjeros. Por esta razón, los hablantes que conocen un idioma con un nivel avanzado tienen más vocabulario virtual en esa lengua que aquellos que hablan dos idiomas con niveles medios o iniciales.

En cuanto a los resultados en lengua española, se produce una influencia directa del nivel de competencia lingüística en lengua extranjera y el número de idiomas. En todos los casos la media observada es mayor conforme asciende el número de idiomas que conoce el informante: 0,45 en informantes que conocen un idioma, 0,47 para la media de dos idiomas, 0,51 para tres idiomas y 0,54 para cuatro idiomas.

Los análisis exploratorios muestran que los resultados en lengua española son bastante más homogéneos tanto entre cada variante como de forma global ya que la incidencia de esta variable es previsiblemente menor que en lengua extranjera. La media entre los cuatro supuestos oscila entre 0,45 y 0,54 mientras que en lengua extranjera era de 0,41 y 0,59. A excepción de la muestra de informantes que conocen cuatro idiomas, el resto de resultados muestra una distribución mesocúrtica tal y como se puede constatar en los datos referidos a la curtosis y al rango.

Dominio de lenguas extranjeras		Estadístico	Error típ.
Una lengua extranjera	Media	0,45055328	0,009888941
	Mediana	0,44870039	

	Desv. típ.	0,079111525	
	Rango	0,343105	
	Curtosis	-0,455	0,59
Dos lenguas extranjeras	Media	0,47324174	0,011886431
	Mediana	0,466191	
	Desv. típ.	0,088152132	
	Rango	0,38547	
	Curtosis	0,313	0,634
Tres lenguas extranjera	Media	0,51420556	0,014215242
	Mediana	0,5218235	
	Desv. típ.	0,087628639	
	Rango	0,343941	
	Curtosis	-0,105	0,75
Cuatro lenguas extranjeras	Media	0,53946337	0,025624239
	Mediana	0,52229183	
	Desv. típ.	0,095877122	
	Rango	0,300787	
	Curtosis	-0,98	1,154

Tabla 5.16. Análisis exploratorios del del conocimiento de lenguas extranjeras con respecto al ICoDD en español

En consonancia con los resultados anteriores, el diagrama de cajas presenta datos homogéneos con un solo valor atípico correspondiente al informante 34. Se trata de un estudiante de la especialidad de Lengua y Literatura que es licenciado en Filología Hispánica y Periodismo por lo que su formación en lengua española se destaca del resto de estudiantes de esa especialidad. La caja más alargada corresponde a los sujetos que conocen cuatro o más idiomas y presenta distribución de los datos leptocúrtica con una curtosis de -0,98.

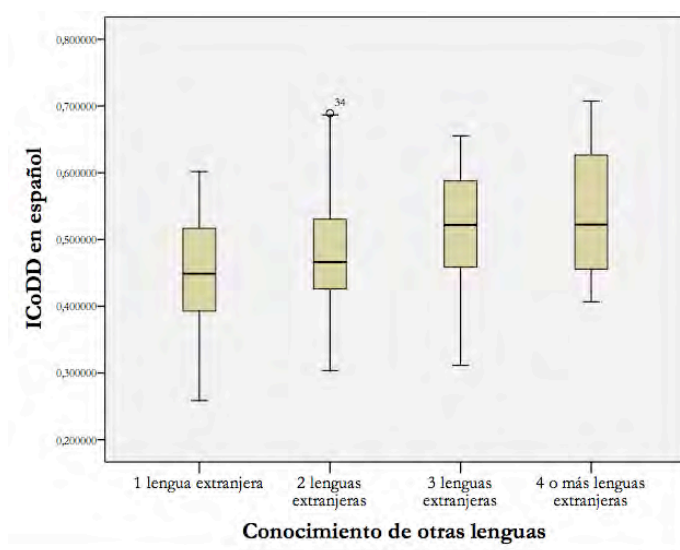


Gráfico 5.21. Diagrama de cajas del conocimiento de otras lenguas con respecto al ICoDD en lengua española

Los análisis exploratorios apuntan que existe una relación significativa entre el conocimiento de idiomas y el léxico de los informantes. Concretamente, se analizó la distribución de la variable del conocimiento de otros idiomas con respecto a la media del ICoDD para cada valor. En un primer análisis, se comprueba que a medida que el informante conocía más de un idioma la media aumentaba. En un análisis posterior, se matiza que el promedio del índice corregido de descentralización léxica en lengua extranjera aumenta en ocasiones más según el dominio de los idiomas que del número de idiomas que habla un informante. Así, los informantes que conocen un idioma con un nivel avanzado tienen mayor caudal léxico que los hablantes que conocen un idioma con un nivel medio y otro con un nivel inicial.

El Gráfico 5.22 muestra las medias del caudal léxico en español y en lengua extranjera según el número de idiomas extranjeros que habla el informante. En ambos casos, la tendencia de la línea es ascendente. En español la diagonal está menos inclinada que en lengua extranjera donde se produce un aumento significativo en el intervalo de los hablantes que conocen dos y tres lenguas²⁴⁴.

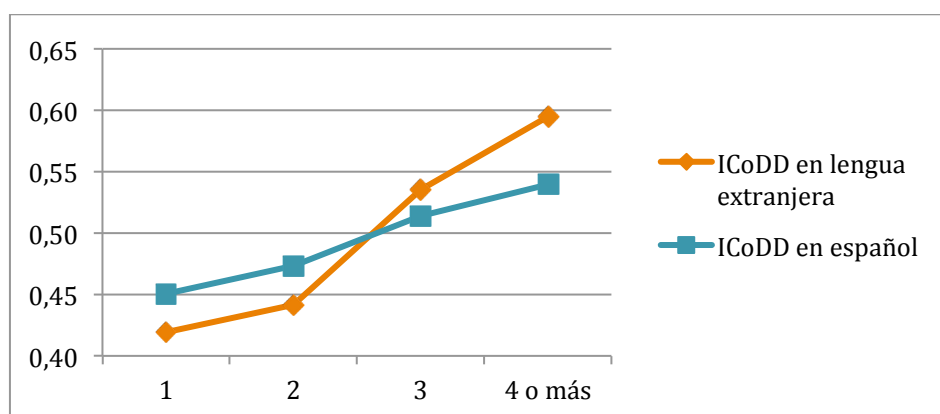


Gráfico 5.22. Distribución del ICoDD en español y en lengua extranjera con respecto al número de idiomas que conoce el informante

Los datos del léxico especializado del Gráfico 5.23 muestran dos diferencias notables. En primer lugar, los resultados de la prueba en español son considerablemente mejores que en lengua extranjera por lo que las líneas no se acercan en ningún momento. En segundo lugar, la tendencia en español y en lengua extranjera es diferente: en la lengua materna la media primero desciende, luego

²⁴⁴ El objetivo del gráfico es mostrar la tendencia de la variable del conocimiento de idiomas en el caudal léxico en español y en lengua extranjera. Debido a que en el cálculo del ICoDD influye el número de palabras e informantes no se compara la media para cada opción.

asciende y por último vuelve a descender mostrando que aparentemente no existe ninguna correlación entre las variables. En cambio, en lengua extranjera la línea es ascendente y se produce un aumento considerable sobre todo en los informantes que hablan uno y dos idiomas extranjeros.

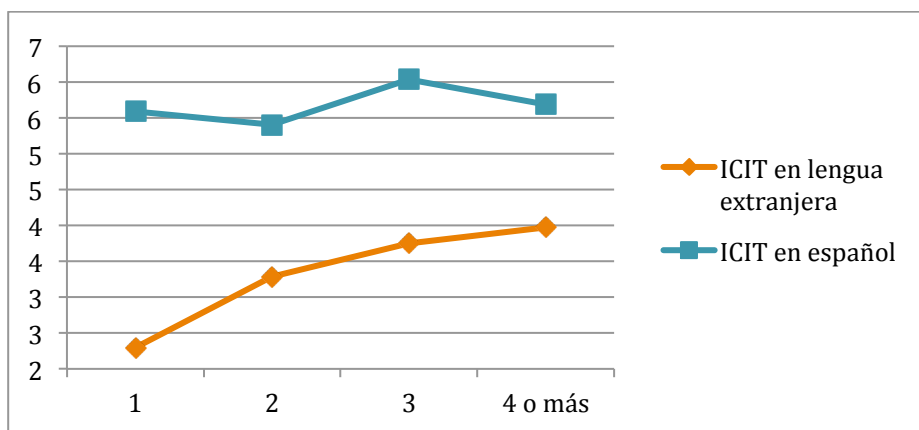


Gráfico 5.23. Distribución del ICIT en español y en lengua extranjera con respecto al número de idiomas que conoce el informante

A continuación, se ha realizado la prueba de Tukey y Scheffé para saber si existen subconjuntos homogéneos según el nivel concreto y el número de idiomas que conoce un informante. Sin embargo, no se han creado subgrupos de forma clara por lo que la opción escogida para los análisis ulteriores es el número de idiomas. Para mostrar estadísticamente la diferencia entre medias de los grupos de esta variable con respecto a las cuatro variables lingüísticas, se ha realizado la prueba de ANOVA de un factor.

El estadístico F es el resultado del cociente entre dos estimadores de la varianza poblacional: el estimador inter-grupos, que es el resultado de las variaciones dentro de las medias de los grupos (en este caso cuatro grupos), y el estimador intra-grupos que, como su nombre indica, se obtiene entre las variaciones dentro de cada grupo. En las variables referidas al caudal léxico es diferente a uno y cuenta con un valor crítico o nivel de significación igual a cero. Por tanto, se rechaza la hipótesis nula de igualdad de medias. Cabe precisar que el valor de F en lengua extranjera es mayor (16,714) que en lengua española (6,944) lo que muestra que los grupos poblacionales según el número de informantes están mejor definidos en inglés y en francés.

En las variables referidas al léxico especializado, los resultados en español y en lengua extranjera son diferentes. En lengua extranjera se rechaza la hipótesis nula porque claramente existen diferencias entre las poblaciones de la muestra ya que el valor F es de 6,944 con un nivel de significación de 0,000. En cambio, en lengua materna el cociente F es prácticamente igual a uno, concretamente 0,925, a un nivel de significación de 0,430 por lo que no hay razón para rechazar la hipótesis nula ya que los cuatro grupos pertenecen a la misma población.

			Suma de cuadrados	gl	Media cuadrática	F	Sig.
ICoDD lengua extranjera	en	Inter-grupos	0,587	3	0,196	16,714	0
		Intra-grupos	1,955	167	0,012		
		Total	2,542	170			
ICoDD español	en	Inter-grupos	0,152	3	0,051	6,944	0
		Intra-grupos	1,218	167	0,007		
		Total	1,369	170			
ICIT lengua extranjera	en	Inter-grupos	69,82	3	23,273	7,396	0
		Intra-grupos	525,503	167	3,147		
		Total	595,323	170			
ICIT español	en	Inter-grupos	9,162	3	3,054	0,926	0,43
		Intra-grupos	550,892	167	3,299		
		Total	560,053	170			

Tabla 5.17. Análisis de ANOVA de un factor de la variable relativa al conocimiento de otros idiomas

Para conocer la relación lineal entre las variables lingüísticas y el conocimiento de otros idiomas se ha calculado el coeficiente de correlación de Pearson. En el caudal léxico en lengua extranjera y en lengua española la relación lineal es significativa. El coeficiente de Pearson para un nivel de significación de 0,001 es de 0,440 en el caso de lengua extranjera mientras que es de 0,331 en lengua española. En cuanto al léxico especializado, la relación solo es lineal en lengua extranjera con un coeficiente de 0,337 y una significación menor de 0,001. En lengua española, la correlación no es significativa: el coeficiente es de 0,67 y la significación es 0,381. Se confirma una vez más que las variables influyen en mayor medida en el léxico activo y sobre todo en lengua extranjera. En español, solo se presentan resultados significativos en el caudal léxico pero carecen de interés en el léxico especializado.

5.4.3.3. Uso de la lengua extranjera en diferentes contextos

Numerosos estudios avalan la efectividad de las estancias en el extranjero en el proceso de adquisición de una lengua (Lennon, 1990; Lapkin, Hart y Swain, 1995; Pérez Vidal y Juan Garau, 2009; Llanes Baró y Muñoz Lahoz, 2009; Llanes Baró y Serrano Serrano, 2011; Llanes Baró, 2011). Los periodos en otro país implican el uso de la lengua en un contexto natural, con necesidades comunicativas reales y donde es frecuente que se forjen relaciones personales, laborales o académicas con otras personas con las que se utiliza la lengua extranjera en el intercambio lingüístico. Por ello, se podría considerar como un factor importante de la variable que engloba el uso de la lengua en diferentes contextos y cuyo cálculo es fruto del análisis factorial del apartado 5.5.1.3.

El cuestionario recoge las diferentes estancias realizadas por los informantes distinguiendo tres tipos: beca Erasmus, estudios becados, estudios sin beca. De esta forma, se relacionan estudios claves como las estancias Erasmus con el léxico disponible de los informantes. Además, se han considerado los motivos de la realización y no realización de la Erasmus, como pueden ser las dificultades económicas o la escasa convalidación de estudios.

a) Beca Erasmus

El número de informantes que ha disfrutado del programa Erasmus asciende a 49, lo que corresponde a un 28,65 % de la muestra total. En los resultados estadísticos se constata que los estudiantes que realizaron la Erasmus presentan una media del ICoDD en lengua extranjera bastante superior (0,54) con respecto a la media de aquellos que no hicieron una estancia en el extranjero con ese programa (0,44).

Los datos referidos a la curtosis, es decir, al grado de concentración que presentan los valores en la región central de la distribución, presentan una distribución normal o mesocúrtica en ambos casos. Por un lado, la curtosis de los estudiantes que realizaron la Erasmus es de -0,025 y, por otro lado, la curtosis de los que no disfrutaron de la beca es de 0,069. En cuanto al rango, es menor en los estudiantes que no han realizado la Erasmus ya que han disfrutado de una beca que

comporta un aprendizaje y uso de la lengua extranjera en situación durante un tiempo prolongado que homogeniza en cierto modo las respuestas.

Realización de la Erasmus		Estadístico	Error típ.
Sí	Media	0,54150849	0,016837795
	Mediana	0,540859	
	Desv. típ.	0,117864567	
	Rango	0,478336	
	Amplitud intercuartil	0,149637	
	Asimetría	-0,437	0,340
	Curtosis	-0,025	0,668
No	Media	0,4365572	0,01004587
	Mediana	0,4306165	
	Desv. típ.	0,110960258	
	Rango	0,554014	
	Amplitud intercuartil	0,162187	
	Asimetría	0,397	0,219
	Curtosis	0,069	0,435

Tabla 5.18. Análisis exploratorios de la realización de la Erasmus con respecto al ICoDD en lengua extranjera

En cuanto a los casos extremos de informantes, se sitúan en la respuesta que indica que no han realizado la Erasmus pero su caudal léxico es muy amplio. Se verifica que se trata de estudiantes que han acabado la licenciatura de Traducción e Interpretación y que, además, han completado afirmativamente otros apartados referidos a estancias en el extranjero y al estudio de lenguas extranjeras en la universidad. Asimismo, el diagrama de cajas muestra que la media de los estudiantes que no han realizado la Erasmus es más simétrica pero los datos son algo más heterogéneos.

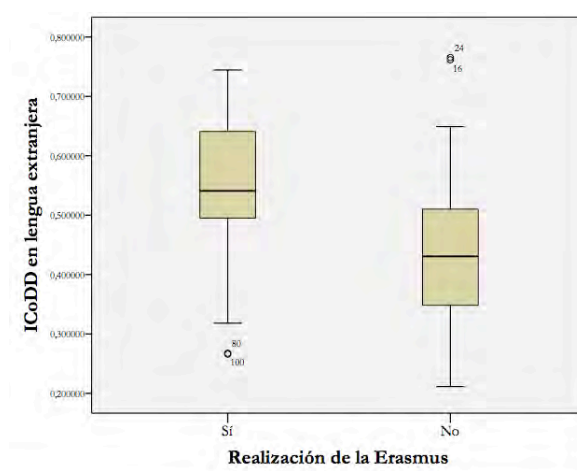


Gráfico 5.24. Diagrama de cajas de la realización de la Erasmus con respecto al ICoDD en lengua extranjera

b) Otros estudios becados en el extranjero

Al igual que ocurre con los resultados anteriores, la estancia en el extranjero influye positivamente en el caudal léxico en lengua extranjera. Los estudiantes que han disfrutado de una beca en el extranjero han sido 74 y la media del ICoDD es de 0,51. Aquellos que no han realizado estancias en el extranjero becas sode becas en el extranjero han sido 97 y la media del ICoDD desciende a 0,43. A diferencia del apartado anterior, en este caso las becas hacen referencia a cualquier tipo de estancia en el extranjero sin menoscabo del tiempo de permanencia en otro país que en muchas ocasiones es de tan solo un mes. Por esta razón, la capacidad léxica de los informantes que han realizado la Erasmus es ligeramente superior de la de informantes que simplemente han disfrutado de una beca en el extranjero.

Realización de una estancia becada en el extranjero		Estadístico	Error típ.
Sí	Media	0,50771655	0,014711814
	Mediana	0,5145585	
	Desv. típ.	0,126555811	
	Amplitud intercuartil	0,160524	
	Asimetría	-0,209	0,279
	Curtosis	-0,28	0,552
No	Media	0,43528731	0,011123787
	Mediana	0,424914	
	Desv. típ.	0,109556596	
	Amplitud intercuartil	0,159211	
	Asimetría	0,456	0,245
	Curtosis	-0,07	0,485

Tabla 5.19. Datos exploratorios de la realización de una estancia becada en el extranjero y el ICoDD en lengua

En cuanto a la distribución de los datos, ambas respuestas presentan una curtosis mesocúrtica (-0,28 y -0,07 respectivamente). La amplitud intercuartil es prácticamente la misma en las dos variantes aunque los datos se muestran más compactos en el caso de los sujetos que han estudiado en el extranjero con algún tipo de beca. El valor atípico corresponde a un informante que no ha disfrutado de ninguna estancia becada en el extranjero pero ha estudiado en el extranjero sin beca, de ahí que el ICoDD sea superior en este gráfico y no así en el Gráfico 5.26.

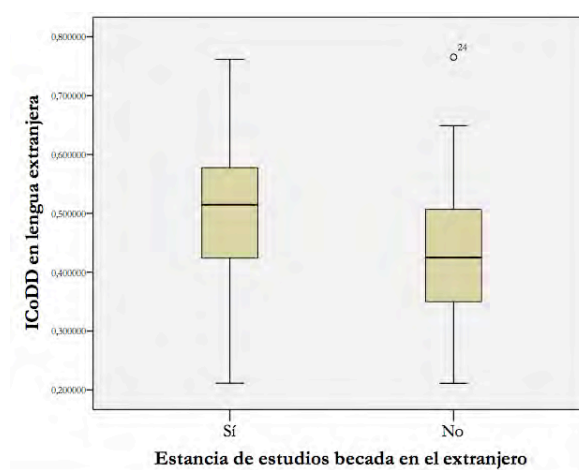


Gráfico 5.25. Diagrama de cajas de la realización de una estancia becada en el extranjero y el ICoDD en lengua extranjera

c) Estudios en el extranjero sin beca

Existen 50 estudiantes que han realizado estancias en el extranjero sin beca frente a los 74 que lo hicieron con algún tipo de beca, por lo que es evidente el papel crucial que juegan los estudios becados y que la capacidad económica de los estudiantes de postgrado es media-alta. El promedio del ICoDD en lengua extranjera de los informantes que han estudiado en el extranjero es de 0,50 y la de aquellos que no lo han hecho es de 0,45. En el primer caso, los datos son más homogéneos que en el segundo, como muestra la amplitud intercuartil (0,15) frente a (0,19), aunque en ambos grupos los datos se concentran en torno a la media.

Realización de estudios no becados en el extranjero		Estadístico	Error típ.
Sí	Media	0,5050717	0,01595175
	Mediana	0,5042775	
	Desv. típ.	0,112795906	
	Rango	0,451975	
	Amplitud intercuartil	0,146028	
	Asimetría	0,327	0,337
	Curtosis	-0,461	0,662
No	Media	0,45074636	0,011176706
	Mediana	0,446615	
	Desv. típ.	0,122943764	
	Rango	0,550354	
	Amplitud intercuartil	0,189448	
	Asimetría	0,248	0,22
	Curtosis	-0,475	0,437

Tabla 5.20. Datos exploratorios de la realización de una estancia sin beca en el extranjero y el ICoDD en lengua extranjera

En los dos apartados anteriores, los casos extremos correspondían a los informantes 16 y 24. No obstante, el informante 24 no sobresale en el diagrama de cajas debido a que forma parte del grupo que ha realizado estudios en el extranjero no becados. Por tanto, el único individuo que sin haber realizado ningún tipo de estancia en el extranjero cuenta con un caudal léxico disponible muy abundante es un estudiante de Traducción e Interpretación que además de los estudios universitarios, ha estudiado idiomas en enseñanzas no regladas.

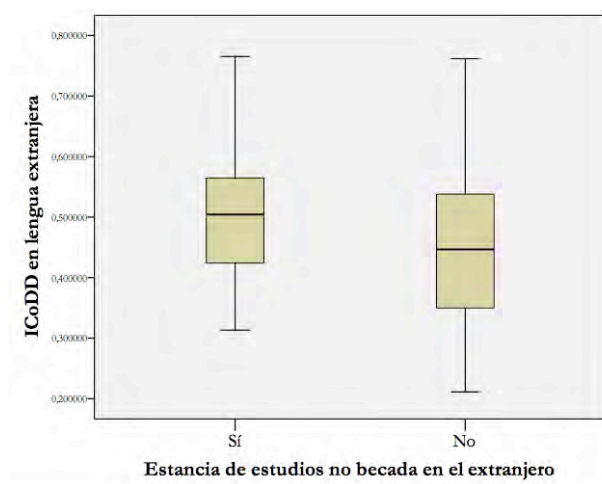


Gráfico 5.26. Diagrama de cajas de la realización de una estancia sin beca en el extranjero y el ICoDD en lengua extranjera

La productividad léxica mostrada por los tres tipos de estancias clasifica la beca Erasmus en un primer rango, con una media de ICoDD de 0,54. A continuación, le sigue la variable relacionada con otros estudios becados con una media de 0,51 y los estudios sin beca con un promedio de 0,50. La media de las palabras aportadas por los informantes que han realizado la Erasmus es 158,18; la de aquellos que han disfrutado de otra beca en el extranjero es de 145,78 y, por último, el promedio de los informantes que han realizado estancias en el extranjero sin beca es de 144,70 palabras.

5.4.3.4. Hábitos de lectura y exposición a los medios de comunicación en español

La hipótesis de partida en este análisis postula que a través de la lectura y los programas culturales el informante puede reforzar su caudal léxico y, sobre todo, el vocabulario pasivo. Para comprobar si guardan una relación positiva, se han

relacionado los hábitos de lectura con el léxico especializado en lengua española. En primer lugar, se ha comprobado que la distribución de la muestra es homogénea y que la media del índice de identificación de tecnicismos aumenta en relación con el número de libros que el informante lee al año. El promedio del ICIT de los informantes que no leen ningún libro es el más pequeño (2,87), a continuación le sigue el tramo de los hablantes que leen hasta cinco libros (4,54) y a partir de ahí se observa que los informantes que leen más de cinco libros al año presentan un léxico especializado similar que va desde 5,88 a 6,40.

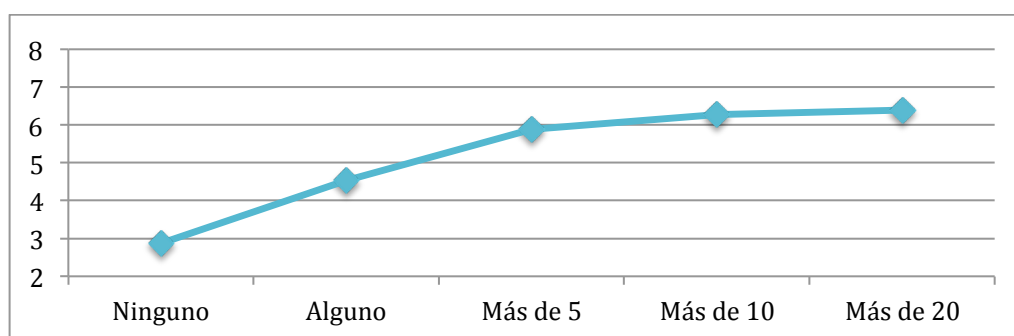


Gráfico 5.27. Promedio del ICIT en español según el número de libros que el informante lee al año

A continuación, se ha calculado la correlación bivariada para conocer si existe una relación lineal entre ambas variables. El coeficiente de Pearson es de 0,162 y el nivel de significancia es menor de 0,05, concretamente 0,035. En cambio, la prueba de la varianza muestra que las poblaciones son muy similares ya que el valor F es de 1,578 a un nivel de significancia de 0,182.

En sintonía con el planteamiento precedente, se ha relacionado la capacidad de identificación del léxico especializado con el número de horas que los estudiantes dedican a escuchar programas culturales e informativos en la radio. Los informantes que consagran menos de dos horas semanales a escuchar programas culturales presentan un promedio del ICIT de 2,97 mientras que aquellos que dedican entre dos y cinco horas han obtenido una puntuación media de 3,39. En este caso, se observa que una gran cantidad de estudiantes no dedica ninguna hora a la semana a escuchar este tipo de programas o lo hace durante menos de dos horas a la semana.

Al unísono de las conclusiones anteriores, se puede confirmar que los programas culturales están relacionados de forma positiva con el léxico especializado ya que el coeficiente de correlación de Pearson es de 0,145 a un nivel de significación de 0,048. La similitud de los datos en las medias conlleva que no se establezcan

diferencias entre muestras poblacionales. Los resultados de estas dos variables ratifican que existen variables que inciden en el léxico especializado de forma leve y que adquieren un papel secundario en el comportamiento de la variable lingüística.

5.4.3.5. Motivación para impartir clases en enseñanza bilingüe

Una de las hipótesis didáctico-críticas que se pretende corroborar es que cuanto mayor sea el caudal léxico en lengua extranjera mayor será la confianza del estudiante para impartir clases en lengua extranjera en un futuro. Además, se considera que existe una relación recíproca entre ambas variables ya que la motivación conlleva que el futuro profesor trate de mejorar su competencia lingüística y, a su vez, el mayor dominio del léxico hace que el hablante adopte actitudes más positivas hacia el aprendizaje y la enseñanza bilingüe. Tras presentar la variable en el apartado 5.3.3.5, se realiza un análisis descriptivo y comparativo de las medias con respecto al caudal léxico en lengua extranjera.

En primer lugar, se destaca que existe cierta progresión en la media del índice corregido de descentralización léxica y la actitud del informante para impartir clases en lengua extranjera²⁴⁵. La media de los informantes a los que les transmite ansiedad es ligeramente superior a los que sienten inquietud. En estos dos primeros grupos se presentan los mínimos más bajos 0,230 y 0,211 respectivamente. A continuación, le sigue la actitud neutra que presenta un notable ascenso en el ICoDD y muestra el valor máximo. En esa misma línea, las actitudes de tranquilidad y placer hacen aumentar el promedio del ICoDD (0,5201 y 0,5412). Los informantes a los que les gustaría impartir clases en lengua extranjera presentan el valor máximo en el caudal léxico (0,7444).

	N	Media	Desv. típica	Error típico	Intervalo de confianza para la media al 95 %		Mín.	Máx.
					Límite inferior	Límite superior		
Ansiedad	26	0,4419	0,1077	0,0211	0,3984	0,4854	0,2300	0,6576
Inquietud	79	0,4252	0,1102	0,0124	0,4005	0,4499	0,2113	0,6448
Actitud neutra	32	0,5158	0,1287	0,0227	0,4694	0,5622	0,3134	0,7653

²⁴⁵ Cabe recordar que esta variable está organizada en forma de escala de Likert en la que la aptitud del hablante se presenta en una escala ascendente.

Tranquilidad	9	0,5201	0,1256	0,0419	0,4236	0,6167	0,2983	0,7370
Placer	25	0,5412	0,1065	0,0213	0,4972	0,5851	0,3847	0,7444
Total	17 1	0,4666	0,1223	0,0094	0,4482	0,4851	0,2113	0,7653

Tabla 5.21. Datos descriptivos de la motivación para impartir clases en lengua extranjera

El gráfico muestra una distribución muy similar de las dos primeras actitudes (ansiedad e inquietud) pero aumenta considerablemente al llegar al grupo central denominado actitud neutra. A partir de ahí la media sube muy ligeramente de tal forma que aparentemente los tres últimos grupos forman parte de un subconjunto poblacional. Por tanto, la motivación podría dividirse en dos grupos para mostrar si el futuro profesorado está motivado o no.

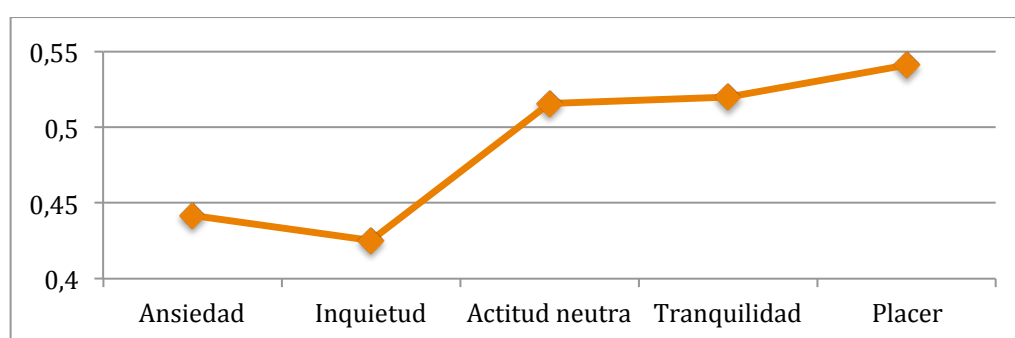


Gráfico 5.28. Distribución del ICoDD en lengua extranjera con respecto a la motivación para impartir clases en lengua extranjera

Por último, se ha realizado la prueba *post hoc* para comprobar la distribución de los subconjuntos homogéneos que aparentemente muestra esta variable. Se presentan los datos de la media armónica de los tamaños de los grupos ya que el número de respuestas para cada opción es diferente. Se constata que tanto por el método de Tukey como de Scheffé se crean los dos subconjuntos comentados anteriormente con una significancia superior al 0,95.

La interpretación de los resultados corrobora que cuanto mayor es el ICoDD en lengua extranjera mejor será la sensación que tiene para impartir clases en esa lengua y viceversa. En clase de lengua extranjera esta información se podría traducir del siguiente modo: si el profesorado cuenta con un alto caudal léxico se sentirá más cómodo para impartir clase en lengua extranjera y, por tanto, empleará más la lengua extranjera que el futuro profesorado que cuenta con un escaso léxico disponible y que le supone ansiedad el hecho de imaginarse impartiendo clases en lengua extranjera.

ICoDD en lengua extranjera				
Sensación al impartir clases en lengua extranjera		N	Subconjunto para alfa = 0,5	
			1	2
HSD de Tukey^{a,b}	Inquietud	79	0,42518203	
	Ansiedad	26	0,44185654	
	Actitud neutra	32		0,51578294
	Tranquilidad	9		0,52014733
	Placer	25		0,54119456
	Sig.		0,989	0,949
Scheffé^{a,b}	Inquietud	79	0,42518203	
	Ansiedad	26	0,44185654	
	Actitud neutra	32		0,51578294
	Tranquilidad	9		0,52014733
	Placer	25		0,54119456
	Sig.		0,994	0,970

Tabla 5.22. Prueba de subconjuntos homogéneos que relaciona la sensación al impartir clases en lengua extranjera y el ICoDD en lengua extranjera

5.5. Análisis multivariante

El análisis bivalente ha mostrado el efecto de algunas variables sobre el vocabulario virtual y la capacidad de identificación de palabras técnicas o especializadas en español y en lengua extranjera, tales como, el número de idiomas que conoce un informante, el contacto con la lengua extranjera en su contexto natural de uso y los estudios previos realizados. En este apartado, se pretende conocer cuáles de esas variables tienen un efecto significativo sobre las variables dependientes cuando se introducen en un modelo multivariante que evalúa, entre otros aspectos, qué efectos son más fuertes y cuáles son redundantes de modo que sea posible relacionar aquel conjunto de variables independientes que se ajustan mejor al comportamiento de la variable dependiente (análisis de regresión lineal múltiple). Además, se realizan análisis de factores para saber si las variables particulares responden superficialmente a dimensiones o variables complejas subyacentes que captan los elementos comunes a todas ellas y permiten afinar la interpretación (análisis multifactorial).

Por último, se representan en un diagrama simple basado en los errores estándar los caminos de influencia entre las variables independientes y entre estas y las variables dependientes, de forma que es posible diseñar modelos, a escala del espacio real, de efectos indirectos y totales entre las variables exógenas, la variable endógena y las variables de mediación entre ambas. El resultado es una representación ideal de fácil captación de las influencias entre variables que subyacen en las listas individuales de la disponibilidad léxica y el léxico especializado en español y en lengua extranjera y sus efectos recíprocos.

5.5.1. Análisis de componentes principales (ACP). Reducción de dimensiones

Las variables presentadas en el apartado 5.3 son una muestra de los datos recopilados en la prueba de disponibilidad léxica, de léxico especializado y del cuestionario sociológico. Dada la importancia concedida en este estudio al campo de la sociolingüística, las respuestas de los informantes han dado lugar a una cantidad ingente de variables que permiten innumerables análisis e interpretaciones. Dentro de ese abanico de variables es posible establecer subgrupos que expliquen el comportamiento de los datos del conjunto que representa.

El procedimiento estadístico aplicado para realizar esa reducción de dimensiones o variables es un tipo de análisis factorial denominado análisis de componentes principales (ACP). Para calcularlo, se introducen las variables originales y se extrae la información común a todas ellas a través de correlaciones múltiples. El resultado consiste en una serie de conglomerados o variables incorreladas ordenadas según la relevancia o variabilidad con la que explican el conjunto. Esos conglomerados o variables resultantes se han incorporado en análisis posteriores de regresión lineal y de modelos causales para dar respuesta a las hipótesis iniciales con una muestra de datos simplificada.

5.5.1.1. Clase social

La clase social se refiere al subgrupo que tiene características socioeconómicas comunes y donde entran en juego criterios económicos y no económicos, tales como

la ocupación y el nivel de instrucción. El cuestionario sociológico incluye preguntas que ayudan medir el estrato social del informante que han sido codificadas para dar lugar a las siguientes variables²⁴⁶:

- El grado de estudios obtenido por los padres²⁴⁷. Cuanto mayor es el nivel de instrucción de los padres mayor es el número asignado.
- Los años concretos de estudio. Este dato viene a reforzar el anterior ya que no solo indica el grado máximo de estudios alcanzado sino también la duración de los mismos. Los años de estudio son sobre todo relevantes en los casos donde los padres han continuado con estudios complementarios que no necesariamente les han permitido obtener un diploma superior, razón por la cual
- La ocupación de los padres. Los informantes aportaron información sobre la profesión concreta que se codificó siguiendo la ordenación del *Real Decreto 1591/2010, de 26 de noviembre, por el que se aprueba la Clasificación Nacional de Ocupaciones 2011*²⁴⁸. En la norma los valores inferiores se refieren a las profesiones más prestigiosas al contrario que en la variable referida al grado de estudios. Por ello, se ha recodificado la ocupación para que conforme aumente el número de la ocupación y de los estudios mejore el estatus social.
- Los ingresos familiares.

El objeto del análisis de componentes es conseguir un conglomerado o conglomerados que ayuden a explicar estas variables en un número más reducido. De esta forma, es posible analizar de forma global la influencia de la clase social de los padres de un individuo en su léxico disponible con una sola variable. Aunque *a priori* la variable relativa a la clase social pudiera ser el resultado del compendio de las variables anteriores, se ha optado por realizar el análisis de componentes principales por dos razones. Por un lado, los resultados estaban marcados sobre todo por la ocupación de los padres ya que va desde cero a once mientras que el resto de variables ofrecen menos variantes. Por otro lado, se pretende obtener un nuevo

²⁴⁶ Corresponden a las preguntas 12, 13, 14 y 15 del cuestionario del Anexo IV.

²⁴⁷ *Vid.* apartado 5.3.2.2 sobre el nivel de estudios.

²⁴⁸ *Vid.* apartado 5.3.2.3 para una información más precisa sobre la codificación.

índice que sea el resultado de análisis estadísticos y no de una mera clasificación realizada a discreción.

La relación entre las variables es apropiada ya que presentan una medida de adecuación muestral de Kaiser-Meyer-Olkin (coeficiente de KMO) de 0,777, el cual indica que el análisis factorial es más adecuado cuanto mayor sea su coeficiente. En cuanto al chi-cuadrado aproximado que resulta de la prueba de esfericidad de Bartlett de es de 774,712 con 21 grados de libertad y una significación de 0. Ello indica que las variables elegidas son adecuadas para un análisis factorial ya que es improbable que la matriz de correlaciones sea una matriz identidad.

KMO y prueba de Bartlett		
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		0,777
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	774,712
	gl	21
	Sig.	0

Tabla 5.23. KMO y prueba de Bartlett del análisis factorial de la clase social

Las comunales muestran que la proporción de las variables explicadas por los factores es mayor en el caso de la educación, seguida de la ocupación y, por último, de los ingresos. La distinción entre las variables según se trate del padre o de la madre es poco significativa. Por ejemplo, en el caso de los estudios del padre la comunalidad adquiere el valor de 0,792 mientras que en el de la madre solo es una centésima inferior 0,782.

Comunalidades	Inicial	Extracción
Estudios del padre	1	0,792
Años de estudio del padre	1	0,856
Ocupación del padre	1	0,533
Ingresos	1	0,34
Estudios de la madre	1	0,782
Años de estudio de la madre	1	0,801
Ocupación de la madre	1	0,567

Tabla 5.24. Comunalidades del análisis de componentes principales de la clase social

Las siete variables se reducen en un solo conglomerado que explica el 66,73 % de la varianza. El resto de componentes presentan unos valores iniciales inferiores a 1, por lo que no se considera significativa su inclusión. Si se realizase el porcentaje de autovariabilidad del resto de variables, se explicaría el 100 % de los valores pero no se

produciría ningún tipo de reducción de factores ya que seguirían siendo siete variables.

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4,671	66,734	66,734	4,671	66,734	66,734
2	0,776	11,08	77,814			
3	0,595	8,499	86,313			
4	0,505	7,21	93,524			
5	0,221	3,159	96,683			
6	0,187	2,678	99,361			
7	0,045	0,639	100			

Tabla 5.25. Prueba de la varianza total explicada del análisis de componentes principales de la clase social

El gráfico de sedimentación o *scree test* confirma los resultados previos y muestra que a partir del segundo componente la línea desciende de forma brusca y se vuelve casi monótona. Por tanto, se podría concluir que este análisis factorial da lugar a una sola variable denominada clase social que abarca aspectos clave para determinar la clase social a la que pertenecen los padres de un individuo, es decir, el nivel de formación alcanzado, los años de estudio, el tipo de ocupación y los ingresos.

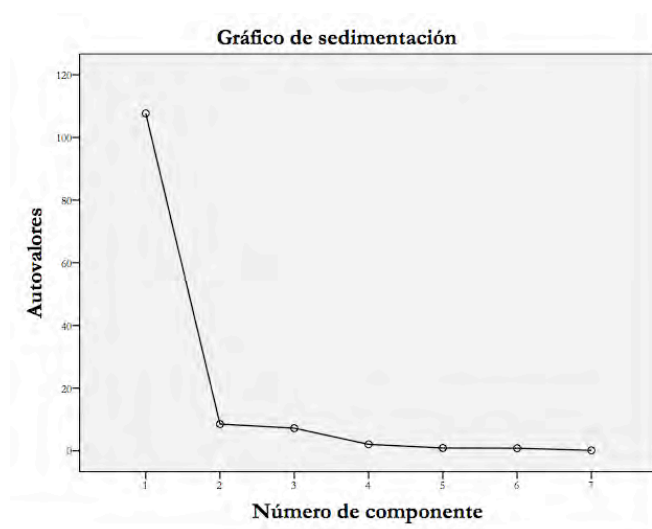


Gráfico 5.29. Gráfico de sedimentación del análisis factorial relacionado con la clase social

5.5.1.2. Motivación para impartir clases bilingües

En este apartado se recogen tres factores relacionados con el interés del futuro profesorado en dar clases en un idioma extranjero, ya sea de un área lingüística

o no lingüística. Esta variable surge de la hipótesis inicial según la cual la motivación de un estudiante influye en su léxico en lengua extranjera. Las características analizadas son las siguientes²⁴⁹:

- La primera variable se centra en la opinión del estudiante sobre su propia formación para impartir clases en un centro bilingüe. Para ello, el sujeto situó en una escala de Likert que va de 0 a 5 el grado de preparación que considera que ha adquirido.
- La segunda variable mide la sensación que le provoca al alumno imaginarse en un futuro impartiendo clases en lengua extranjera. El informante tuvo que elegir entre las siguientes opciones: ansiedad, inquietud, indiferencia y placer. La codificación asignó valores en escala de 0 a 4.
- La tercera variable muestra el interés de los estudiantes en trabajar en un centro bilingüe. Las respuestas podían ser de tres tipos: negativas, dudosas (no sabe/no contesta) o afirmativas. Se codificaron de 0 a 2 respectivamente²⁵⁰.

Los indicadores resultantes muestran un análisis factorial satisfactorio ya que la prueba de KMO explica el modelo en un 0,64 y la prueba de esfericidad de Barlett en un 85,427 %. En cuanto al análisis de la varianza total explicada revela que las tres variables se reducen a un solo componente. Este hecho se constata por los autovalores iniciales y las sumas de las saturaciones al cuadrado de la extracción. En el primer caso, según el criterio de Kaiser en la columna que muestra el total de los autovalores iniciales tan solo se conservan los valores mayores que uno, lo que corresponde en este análisis al primero (1,83). Por tanto, considerando que el primer autovalor es de 1,83 y que existen 3 variables el resultado el porcentaje de autovariabilidad es el 61,245 %.

$$100 \cdot \frac{1,83}{3}$$

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	1,837	61,245	61,245	1,837	61,245	61,245
2	0,678	22,616	83,861			

²⁴⁹ Corresponden a las preguntas 53, 50 y 49 del cuestionario del Anexo IV.

²⁵⁰ En un segundo estadio se codificaron los motivos por los que los informantes estaban interesados o no en impartir clases en un centro bilingüe ya que era una de las opciones de la pregunta 49 del anexo IV.

3	0,484	16,139	100
---	-------	--------	-----

Tabla 5.26. Prueba de la varianza total explicada del análisis de componentes principales de la motivación para impartir clases bilingües

En la matriz de componentes los valores de las tres variables superan el 0,7, donde la variable referida a la opinión sobre la formación para impartir clases en un centro bilingüe obtiene el máximo valor: 0,833. Por tanto, se podría concluir que estos tres factores dan lugar a una variable nueva que presenta la motivación del futuro profesorado para impartir clases bilingües. Aunque la información recopilada corresponde a tres preguntas diferentes la correlación entre sí es significativa tanto estadística como cualitativamente.

Matriz de componentes	Componente
	1
Opinión sobre la propia formación para impartir clases en un centro bilingüe	0,833
Sensación impartiendo clases en lengua extranjera	0,783
Le gustaría trabajar en un centro bilingüe	0,729

Tabla 5.27. Resumen de la matriz de componentes en espacio rotado de la motivación para impartir clases bilingües

5.5.1.3. Uso de la lengua en diferentes ámbitos

Este análisis trata de poner en relación la información recogida por medio de variables que hacen referencia al uso de una lengua extranjera en diferentes contextos. El estudio factorial se lleva a cabo para dar respuesta a una de las hipótesis iniciales que postula que cuanto mayor sea el contacto de un individuo con una lengua extranjera mayor será su competencia lingüística y, por tanto, su léxico disponible. Las variables analizadas son las siguientes:

- Densidad de las relaciones. Este factor es el resultado de un compendio de siete ítems que tratan de situar el grado de concentración de las relaciones de los informantes. El resultado se entiende dentro de una escala que oscila entre un mínimo de densidad marcado por el número 0 y un máximo al que le corresponde el número 6²⁵¹.

²⁵¹ Esta variable corresponde a la pregunta 32 de la investigación de Ávila Muñoz y Villena Ponsoda (2010: 312) y a la pregunta 17 del Anexo IV. Cada respuesta positiva suma un punto por lo que a mayor número de respuestas positivas mayor será el índice de densidad.

- Sociabilidad²⁵². En contraposición con la variable anterior, los informantes tienen que responder si las personas con las que se relacionan están en su barrio, pueblo, en Málaga, en otros lugares de Andalucía, en otras ciudades de España o tanto en España como en el extranjero.
- Realización de una estancia en el extranjero. Se ha incluido el tipo de beca Erasmus por ser prolongada en el tiempo (al menos un cuatrimestre) y por implicar un contacto lingüístico considerable. Las posibles respuestas eran afirmativas (1) o negativas (0). Además, en los análisis bivariantes del apartado 5.4.3.3 se ha mostrado que existe una correlación significativa entre el caudal léxico en lengua extranjera y la realización de una estancia en el extranjero disfrutando de la beca Erasmus.
- Suma del uso de las lenguas en diferentes contextos. En esta variable el informante tuvo que marcar con una cruz el uso de la lengua española, de la lengua extranjera de la prueba y de otra lengua, si la hubiere: en clase, con amigos y con la familia. Además, se incluyó una referencia a la asiduidad de la comunicación que se presenta en una escala de 4 a 0: todos los días, varias veces por semana, con menos frecuencia y nunca. De esta forma, si un estudiante afirma hablar todos los días (4 puntos) con amigos en lengua extranjera y asiste a clases de esa lengua una vez a la semana (2 puntos) se le suman los números correspondientes de cada frecuencia y se obtiene el total del uso de la lengua (6 puntos)²⁵³.
- Medios de comunicación²⁵⁴. Al igual que en el apartado anterior, los sujetos marcaron el uso de la lengua en relación con los medios de comunicación. Se considera que en lengua extranjera con independencia del tipo de programa que vea u oiga el informante, cualquier exposición a los medios de comunicación contribuye en mayor o menor medida a reforzar el conocimiento del idioma.

²⁵² Esta variable aparece en la pregunta 29 del estudio de Ávila Muñoz y Villena Ponsoda (2010: 311) y a la pregunta 16 del Anexo IV. Las respuestas van de 0 hasta 6 que aumentan a medida la red de contactos del informante sea más abierta.

²⁵³ Corresponde a la pregunta 25 del Anexo IV.

²⁵⁴ Corresponde a la pregunta 27 del Anexo IV. La codificación asignó 0 a ninguna hora; 1 a menos de 2 horas; 2 a el tramo entre dos y cinco horas y 3 a más de cinco horas.

La medida de adecuación muestral de Kaiser-Meyer-Olkin indica que el análisis de componentes está explicado en un 0,765 siendo el valor del límite máximo 1. Asimismo, las comunalidades extraídas en todos los casos son superiores al 0,5 y en dos casos superior al 0,9.

Comunalidades	Inicial	Extracción
Sociabilidad	1	0,93
Suma del uso de la lengua extranjera en diferentes contextos	1	0,823
Suma del uso de las lenguas que conoce el informante en diferentes contextos	1	0,797
Medios de comunicación	1	0,578
Erasmus	1	0,527
Densidad de las relaciones	1	0,927

Tabla 5.28. Comunalidades del análisis de componentes principales del uso de la lengua

El análisis resultante reagrupa las variables en dos conglomerados que explican un total de la varianza del 76,375 %. El primer componente lo forman las variables relacionadas con el uso de la lengua en diferentes contextos (suma de la lengua extranjera según grados, suma total uso de las lenguas, realización de la Erasmus y exposición a los medios de comunicación). El segundo componente relaciona las dos variables que miden el grado de relación de un informante con otras personas: la densidad y la sociabilidad.

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acum.	Total	% de la varianza	% acum.	Total	% de la varianza	% acum.
1	3,284	54,727	54,727	3,284	54,727	54,727	2,655	44,249	44,249
2	1,299	21,648	76,375	1,299	21,648	76,375	1,928	32,127	76,375
3	0,573	9,557	85,933						
4	0,521	8,679	94,612						
5	0,192	3,202	97,814						
6	0,131	2,186	100						

Tabla 5.29. Comunalidades del análisis de componentes principales del uso de la lengua

A continuación, se muestran los datos que corresponden a combinaciones lineales de las variables originales para los componentes rotados. En el componente uno se observa que los datos destacados en negrita tienen un peso mayor que el resto. Todos estas variables tienen en común el uso de la lengua por parte del informante: en diferentes contextos, a través de la exposición a los medios de

comunicación y en un contexto natural a través de las situaciones de comunicación en una estancia en el extranjero.

En cambio, en el segundo componente las variables con valores entorno a más o menos uno polarizan sus resultados: el tratamiento con personas pasa a tener un valor de 0,941 y la densidad de -0,944. Estas cifras opuestas forman un conglomerado que podría denominarse *Red social* y que se refiere al tamaño de la red de vínculos entre el informante y otros hablantes. A su vez, los factores con mayor carga en el primer componente presentan un valor considerablemente inferior en el componente dos lo que ratifica la interpretación dada para este análisis. Cabe precisar que la beca Erasmus presenta el nivel más alto en este segundo conglomerado lo que lleva a pensar que además de evaluar el uso de la lengua también influye en las características de las redes sociales que crea el informante.

Matriz de componentes rotados ²⁵⁵	Componente	
	1	2
Suma del uso de la lengua extranjera de la prueba en diferentes contextos	0,904	0,073
Suma de otra lengua extranjera que conoce el informante en diferentes contextos	0,879	0,158
Medios de comunicación	0,726	0,225
Beca Erasmus	0,676	0,263
Tratamiento con personas	0,209	0,941
Densidad	-0,191	-0,944
Método de rotación: Normalización Varimax con Kaiser.		
La rotación ha convergido en 3 iteraciones.		

Tabla 5.30. Prueba de la varianza total explicada del análisis de componentes principales del uso de la lengua

La hipótesis de partida pretende corroborar que los informantes con un mayor uso de la lengua extranjera disponen de un léxico más amplio que el resto. Asimismo, ese léxico tiene una relación indirecta con el tipo de contactos de los hablantes a través del uso de la lengua ya que los sujetos con redes más amplias, sobre todo fuera de España, tendrán más ocasiones de utilizar la lengua extranjera. Los componentes creados serán objeto de análisis multivariantes posteriores para simplificar los resultados. No obstante, en ocasiones se ha indagado con algunas de las variables individuales para determinar cuál de esas características concretas llevan a un estudiante a proporcionar un mejor caudal léxico.

²⁵⁵ La matriz de componentes rotados muestra que la rotación ha convergido en tres iteraciones.

5.5.2. Análisis de regresión lineal

El análisis de regresión múltiple explica la relación que ejerce un conjunto de variables independientes (X_i) sobre una variable dependiente (Y) (Calvo Gómez, 1993: 193). El objetivo de este modelo estadístico es doble: por una parte predecir el funcional que relaciona las variables y, por otra parte, comprender las interrelaciones entre las mismas. Según Calvo Gómez (1993: 194):

El principio metodológico-estadístico de la Regresión Múltiple consiste en encontrar aquella función que mejor prediga Y , es decir, aquella que a partir de las variables independientes X menor error cuadrático de predicción cometa.

Para Albarrán Lozano, Guillén Estany y Ayuso Gutiérrez (2003: 279-280) el modelo de regresión es concebido como la antesala de la interpretación de los resultados de investigaciones empíricas. El modelo de regresión lineal múltiple se rige en función del concepto de dependencia estocástica entre las variables. Por ello, el modelo lleva incluido un término de perturbación aleatoria que intenta corregir ese valor azaroso.

En esta sección se lleva a cabo un análisis de modelación lineal automatizada con el objetivo de crear un modelo estándar único que explique las relaciones entre campos. El modelo vincula la distribución aleatoria del objetivo o variable endógena con diferentes predictores lineales extraídos de las variables exógenas. En este tipo de análisis se incluyen todas las variables que cumplen determinados requisitos ya que el resto de variables se desestiman y no influyen en los resultados. El método seleccionado es el método *forward* o paso adelante en el que los regresores correlacionados significativamente con la variable dependiente se introducen uno a uno.

A partir de este patrón, se realizan cuatro tipos de análisis de regresión lineal cuyo objetivos se conectan directamente con las hipótesis iniciales. El primer y segundo análisis sirven para determinar la relación multicausal del caudal léxico en lengua española y en lengua extranjera con respecto a variables independientes o regresoras. El tercer y cuarto análisis se centran en explicar la influencia de variables exógenas en la capacidad de identificar palabras especializadas en español y en lengua

extranjera. De esta forma, se pretende conocer cuáles son los factores que inciden en el vocabulario activo y pasivo en lengua materna y lengua extranjera.

5.5.2.1. Análisis de regresión lineal del caudal léxico en lengua extranjera

En este apartado se ha realizado un análisis de modelo lineal con objeto de predecir un objetivo continuo en base a relaciones lineales entre el objetivo, el índice corregido de descentralización léxica en lengua extranjera, y los predictores o entradas a través de las cuales se pretende corroborar o refutar algunas de las hipótesis iniciales. Los predictores que se considera que tienen un efecto sobre el caudal léxico en lengua extranjera son el conocimiento de otros idiomas, la especialidad del máster, la motivación del profesorado para impartir clases bilingües y el caudal léxico en lengua española. La relación bivariante de estos factores con respecto a la variable dependiente, es decir, el índice corregido de descentralización léxica se ha comprobado previamente que es significativa en el apartado anterior pero falta por conocer en qué medida influyen en este modelo multivariante. Asimismo, se ha incluido la variable referida a la clase social resultante del análisis de componentes principales pero se espera que su influencia sea secundaria ya que se comprobó que la formación del padre era poco relevante.

El método de selección de modelos utilizado es el denominado paso adelante. En un primer momento se introdujeron todas variables que podrían tener mayor efecto como una mayor cantidad de variables pero gracias a este método solo se han seleccionado aquellas variables relacionadas significativamente con el ICoDD en lengua extranjera. El modelo se construye utilizando el criterio de información que mide el grado de calidad de la regresión múltiple según el subconjunto de variables. Concretamente se ha presentado el criterio de información Akaike corregido el cual determina el grado de información perdida cuando se conmutan un conjunto de datos reales por un modelo más sencillo. El ajuste del modelo será más adecuado cuanto menor sea su valor y en cada paso aumenta ligeramente ese valor partiendo del paso uno donde el criterio era de -8,38,819 hasta el quinto paso donde toma el valor de -893,462.

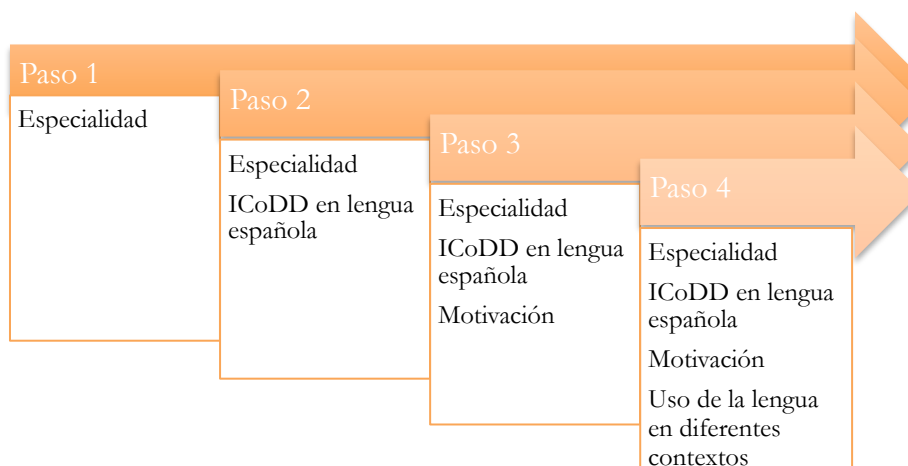


Gráfico 5.30. Pasos del modelo de regresión lineal del ICoDD en lengua extranjera

En cuanto al nivel de precisión, calcula el rigor de predicción y estimación del modelo. Cuanto más se acerca al 100 % más preciso es el modelo. En el primer paso el nivel de precisión es del 52 % pero al introducir el resto de variables el modelo asciende al 74,3 %. Asimismo, el modelo incluye la preparación automática de datos utilizada en el análisis y detección de problemas que filtra campos conflictivos, corrige errores y crea nuevos atributos en los casos pertinentes. Estos cambios son analizados con más detenimiento ya que se refieren sobre todo a la recodificación de variables.

El pronóstico por observado muestra en un diagrama de dispersión los valores pronosticados del eje de ordenadas en el eje de abscisas. Los puntos deben situarse en una línea de 45 grados para que el pronóstico sea apropiado. En este caso se observa que hay ciertos registros que se desplazan levemente de la línea ideal pero sin llegar a mostrar un pronóstico inadecuado.

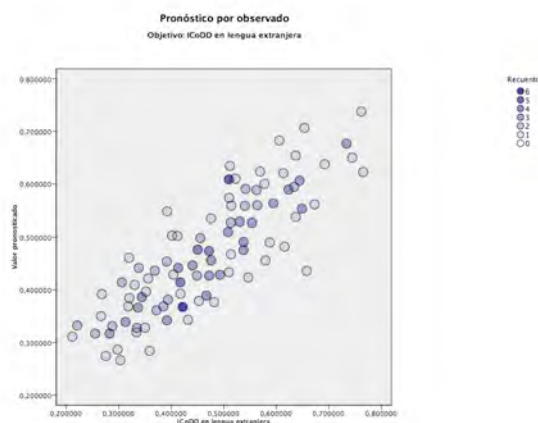


Gráfico 5.31. Pronóstico por observado del ICoDD en lengua extranjera

En cuanto al histograma de los residuos estudentizados, compara la distribución de los residuos con una distribución normal. Para interpretarlo, es necesario observar la distancia de las frecuencias de los residuos con respecto a la línea que representa la distribución normal. Por tanto, cuanto más se acerquen ambos parámetros más normalizada será la distribución de los residuos. En líneas generales, las frecuencias del gráfico Gráfico 5.32 se acercan a la línea de la distribución normal ya que se sitúan dentro del rango de ± 3 .

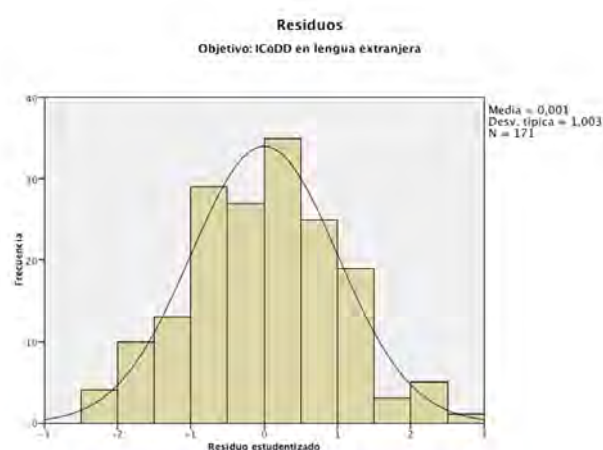


Gráfico 5.32. Histograma de residuos estudentizados del ICoDD en lengua extranjera

A continuación, se presentan en la Tabla 5.31 los datos de la estimación del coeficiente que representa el tipo de efecto de cada variable sobre el ICoDD. En primer lugar, se observa que el coeficiente Beta de la intersección o constante es de 0,389. A continuación le sigue el coeficiente para el resto de variables y sirve para conocer cuanto aumenta la variable dependiente cuando sube un punto la variable independiente. Las variables independientes están ordenadas según la importancia en el modelo de forma que cuanto más se acerquen a 1, que es el valor total, mayor será el peso en el modelo.

La variable independiente que aparece en el primer paso y con mayor importancia es la especialidad del máster. Se observa que existen cinco grupos de especialidades del máster de las cuales algunas inciden positivamente en el ICoDD pero el resto hace disminuir ese coeficiente. Cabe destacar que este método permite incluir variables categóricas o cualitativas, tales como la especialidad del máster, a través de variables ficticias (*dummy variables*). Este tipo de variables, también conocidas como variables binarias o indicadoras, tienen un valor entre 0 y 1 y sirven para

representar un factor cualitativo. Ello permite que los análisis de regresión presentados no solo se limiten a variables cuantitativas.

También se observa que el ICoDD en español aumenta el ICoDD en lengua extranjera. En cambio, las variables que hacen referencia al conocimiento de otras lenguas y la motivación del informante hacen disminuir el ICoDD en lengua extranjera, cuestión algo contradictoria que se examina con más detalle a continuación. Por último, la clase social cobra poca importancia en el modelo 0,029 y el coeficiente beta es de tan solo 0,011 a un nivel de significación de 0,38.

	Coeficiente Beta	Error típico	t	Significación	Nivel de confianza		Importancia
					Inferior	Superior	
Intersección	0,203	0,039	7,583	0,000	0,217	0,370	
Especialidad. Grupo 1	0,069	0,021	3,230	0,001	0,027	0,112	0,494
Especialidad. Grupo 2	-0,082	0,026	-3,097	0,002	-0,134	-0,030	0,494
Especialidad. Grupo 3	-0,077	0,014	-5,237	0,000	-0,106	-0,048	0,494
Especialidad. Grupo 4	0,016	0,018	1,860	0,391	-0,021	0,052	0,494
Especialidad. Grupo 5	0,000 _a ²⁵⁶						0,494
ICoDD en español	0,465	0,074	6,313	0,000	0,320	0,611	0,317
Motivación	-0,050	0,012	-4,091	0,009	-0,075	-0,026	0,133
Motivación 2	0,000 _a						
Uso de la lengua en diferentes contextos	0,018	0,007	2,651	0,009	0,005	0,032	0,056

Tabla 5.31. Datos estadísticos sobre el modelo de regresión lineal del caudal léxico en lengua extranjera

5.5.2.1.1. Análisis de medias estimadas de los predictores

a) Especialidad del máster

El gráfico de medias estimadas presenta el efecto de cada tipo de especialidad en el caudal léxico en lengua extranjera. En primer lugar, la especialidad de Lengua Extranjera Inglés se sitúa con el nivel superior con una media de ICoDD de 0,583 seguida por el último conjunto de variables compuesto por las especialidades de Lengua Extranjera Francés y Lengua Castellana y Literatura con un promedio de 0,528. Superando la línea de 0,5 se encuentran las especialidades de Geografía e Historia, Dibujo, Música y FOL con 0,504. La media del ICoDD es de 0,467 lo cual

²⁵⁶ Este coeficiente está establecido en cero porque es redundante.

muestra que estos subgrupos de especialidades influyen de manera positiva en el ICoDD en lengua extranjera. En cuanto a las especialidades de Matemáticas, Orientación, Hostelería, Turismo y TIPI se relacionan de forma negativa con el ICoDD mostrando valores por debajo de la media, especialmente el caso de Matemáticas con un 0,421.

En el se comentó que la influencia positiva y negativa de ciertos grupos de especialidades. Concretamente se destaca que la especialidad de Lengua Extranjera Inglés influye de forma positiva y los subgrupos formados por las especialidades de Matemáticas, Orientación, Hostelería, Turismo y Tecnología, Informática y Procesos Industriales lo hacen de forma negativa. La incidencia del grupo formado por las especialidades de Geografía e Historia, Dibujo, Imagen y Artes Plásticas y Música no es relevante, por lo que no se muestra en el diagrama.

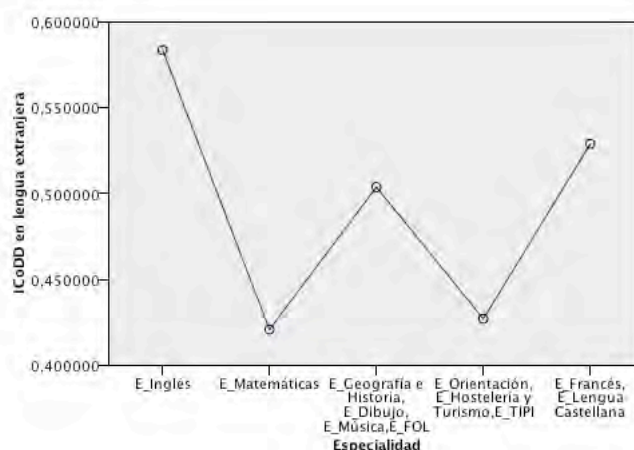


Gráfico 5.33. Medias estimadas del ICoDD en lengua extranjera según especialidades del máster

Cabe destacar que en la codificación originaria cada especialidad del máster tiene atribuido un número de identificación. En cambio, la preparación de datos automática realizada en el cálculo de la regresión lineal presenta las especialidades reagrupadas en cinco subconjuntos. Con objeto de comprobar la homogeneidad de dichos subgrupos se aplicó un análisis de medias, concretamente Anova de un factor seleccionando las pruebas de Tukey y Scheffé. Sendos estudios de rango *post hoc* permiten identificar subconjuntos homogéneos de medias que no se diferencian de forma significativa. Por un lado, el método de la diferencia significativa honesta (*Honestly Significant Difference*) o método de Tukey realiza todas las combinaciones por

pares posibles utilizando el rango estudentizado. Por otro lado, el método Scheffé lleva a cabo comparaciones múltiples por pares con la distribución muestral F .

Para llevar a cabo el análisis de medias se utilizó la variable referida a la especialidad ya recodificada con un valor diferente para cada subgrupo creado de forma automática²⁵⁷. La Tabla 5.32 presenta las medias para los grupos en los subconjuntos homogéneos empleando el tamaño muestral de la media armónica (23,737) que se utiliza porque la medida de los grupos no son iguales, tal y como se muestra en la tercera columna. Se constata un nivel de significancia de 1 en cada subconjunto creado de forma automática en el proceso de regresión. Por tanto, en el caso de la media del ICoDD en lengua extranjera de las diferentes especialidades del máster se crea una variable que distingue entre estos cinco subconjuntos, permitiendo la búsqueda de semejanzas y diferencias entre cada uno de ellos.

ICoDD en lengua extranjera							
Especialidad del máster		N	Subconjunto para alfa = 0,5				
			1	2	3	4	5
HSD de Tukey^{a,b}	2	10	0,3231				
	3	60		0,3879			
	5	44			0,4679		
	4	32				0,5338	
	1	25					0,6250
	Sig.		1,000	1,000	1,000	1,000	1,000
Scheffé^{a,b}	2	10	0,3231				
	3	60		0,3879			
	5	44			0,4679		
	4	32				0,53376	
	1	25					0,6250
	Sig.		1,000	1,000	1,000	1,000	1,000

Tabla 5.32. Prueba de subconjuntos homogéneos que relaciona la especialidad del máster y el ICoDD en lengua extranjera

b) Índice corregido de descentralización léxica en español

El Gráfico 5.34 corrobora el análisis de correlación lineal bivalente del apartado 5.4.1.3 donde se muestra la relación positiva entre el caudal léxico de la lengua materna y la lengua extranjera ya que al aumentar el ICoDD en una lengua también lo hace en la otra aunque no en la misma proporción. Al extrapolar los resultados se podría concluir que los factores extralingüísticos que inciden

²⁵⁷ Grupo 1: Lengua Extranjera Inglés. Grupo 2: Matemáticas. Grupo 3: Geografía e Historia; Dibujo, Imagen y Artes Plásticas; Música; Formación y Orientación Laboral. Grupo 4: Orientación Educativa; Hostelería y Turismo; Tecnología, Informática y Procesos Industriales. Grupo 4: Lengua Extranjera Francés; Lengua Castellana y Literatura.

significativamente en el ICoDD de una lengua lo hacen de forma indirecta en el ICoDD de otra lengua y, por tanto, si la especialidad tiene un papel predominante en el léxico en lengua extranjera también intervendrá en el léxico en español.

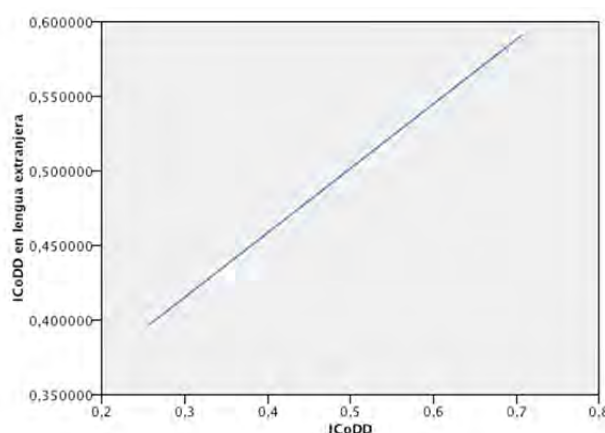


Gráfico 5.34. Medias estimadas del ICoDD en lengua extranjera y el ICoDD en español

c) Motivación para impartir clases en lengua extranjera

La relación entre el caudal léxico en lengua extranjera y la sensación que tiene el informante al impartir clases en esa lengua aparece con una incidencia negativa en el análisis de regresión. La codificación original otorga una puntuación en forma de escala de Likert que aumenta según mejora la seguridad del informante²⁵⁸. No obstante, al realizar el análisis de regresión las distintas respuestas de la variable se han recodificado en dos subconjuntos homogéneos.

- El primero recoge las sensaciones de ansiedad e inquietud para impartir clase en lengua extranjera.
- El segundo agrupa las respuestas referidas a la actitud neutra, la tranquilidad y el placer.

Esta organización de las variantes de esa variable se realizó en el apartado 5.4.3.5 del análisis bivalente²⁵⁹. En este caso se observa que el subconjunto homogéneo referido a las sensaciones de ansiedad e inquietud ejercen una influencia negativa sobre el modelo multivariante. El segundo subgrupo formado por las

²⁵⁸ Se otorgaron los siguientes valores: ansiedad (0), inquietud (1), actitud neutra (2), tranquilidad (3) y placer (4).

²⁵⁹ Para construir este modelo se analizaron por separado todas las variables referidas a la motivación del profesorado para impartir enseñanza bilingüe. Se ha incluido la variable con mayor incidencia para mostrar el factor más relevante en este sentido. En cambio, en el modelo causal se ha optado por incluir el genérico que representa al conjunto de variables que se refieren a la motivación obtenido en el análisis de componentes principales del apartado 5.5.1.2.

sensaciones de actitud neutra, tranquilidad y placer no ejercen ningún tipo de influencia sobre en el modelo de regresión. Ello indica que los estudiantes con menor caudal léxico se muestran más inseguros a la hora de impartir clases en lengua extranjera, hecho que a su vez se puede inferior en sentido inverso, es decir, aquellos sujetos que se muestran menos motivados no muestran tanto interés o firmeza al realizar la prueba de disponibilidad léxica y aportan menos vocablos.

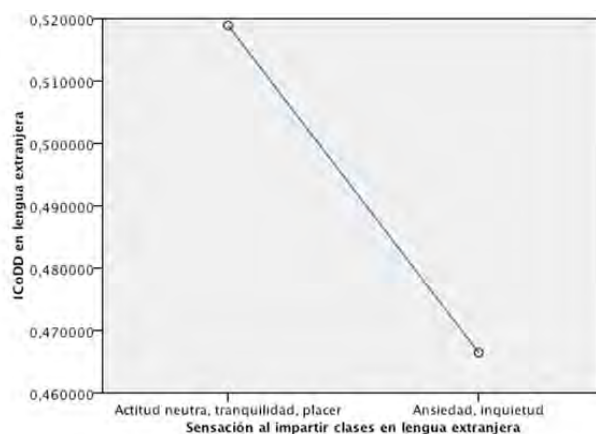


Gráfico 5.35. Medias estimadas entre el ICoDD en lengua extranjera y la sensación al impartir clases en lengua extranjera

d) Uso de la lengua extranjera en diferentes contextos

El análisis de regresión incluye la variable del uso de la lengua extranjera en diferentes contextos calculada en el análisis factorial del apartado 5.5.1.3. La correlación de Pearson a un nivel de significación bilateral de 0,01 es de 0,489 mostrando así una estrecha relación entre el uso de la lengua en diferentes contextos y el caudal léxico en lengua extranjera.

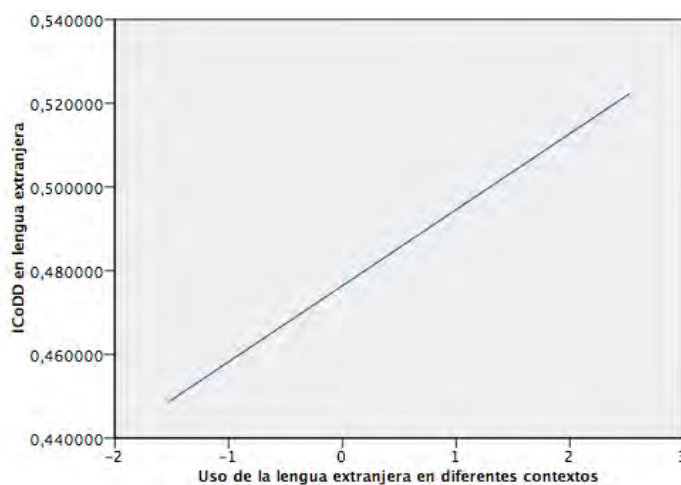


Gráfico 5.36. Medias estimadas del ICoDD en lengua extranjera y el uso de la lengua extranjera en diferentes contextos

Una vez que se ha comprobado la relevancia de esta variable, se ha pretendido localizar qué elementos tienen mayor peso. Se destacan los hábitos de lectura donde es posible diferenciar dos grupos: aquellos informantes que no leen ningún libro al año y los que leen al menos uno. La media armónica calculada es de 0,417 en el caso de los informantes que no leen ningún libro en lengua extranjera al año; 0,502 para los que leen alguno; 0,595 para los que leen entre 5 y 10 libros y 0,616 para los que leen más de 10 libros al año.

Otro de los aspectos que se destaca es la realización de estudios fuera de España donde tiene una influencia mayor la realización de la Erasmus²⁶⁰ y el visionado de películas en lengua extranjera. En esta última variable se han destacado dos subconjuntos homogéneos: aquellos que no ven ninguna película en lengua extranjera y los que ven alguna película semanalmente con independencia de la cantidad. En definitiva, estas variables reiteran la importancia del uso de la lengua en contexto y de la importancia del contacto con la lengua extranjera en el aprendizaje de una lengua.

5.5.2.2. Análisis de regresión lineal del caudal léxico en español

El léxico disponible en lengua española es más homogéneo que en lengua extranjera ya que el nivel de instrucción de los estudiantes es muy similar (diplomatura, ingeniería o licenciatura) y el español es su lengua materna. La formación académica se dibuja como la variable más influyente en los sujetos y el resto de variables tienen una incidencia menor que en lengua extranjera. No obstante, se espera que los hablantes que han realizado estudios ligados a algún área lingüística hayan profundizado en la lengua española y demuestren un léxico más descentralizado. A su vez, se incorporan las variables relacionadas con los hábitos de lectura y el caudal léxico en lengua extranjera por haber mostrado en los análisis bivariantes su incidencia.

Las variables seleccionadas en este modelo han sido: el ICoDD en lengua extranjera, la especialidad del máster, la capacidad de identificación de tecnicismos y los hábitos de lectura. El modelo se explica en un 46,3 % con un criterio de

²⁶⁰ Vid. apartado 5.4.3.3.

información de -925,457 lo que constata la mayor incidencia de las variables sobre el modelo en lengua extranjera. El pronóstico por observado muestra que la distribución de los valores no es del todo adecuada ya que dista de crear una línea de 45 grados. La mayoría de los datos se concentran en torno a la media dibujando una línea ascendente poco inclinada.

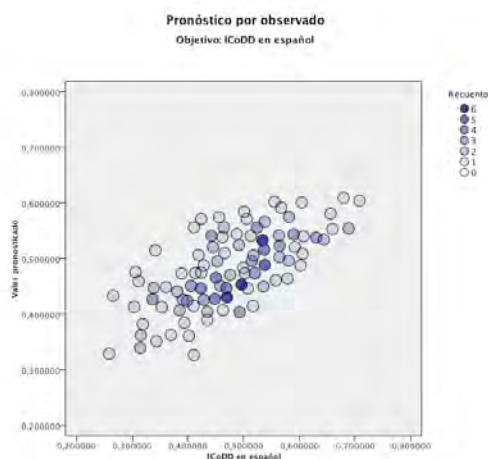


Gráfico 5.37. Pronóstico por observado del ICoDD en lengua extranjera

El histograma de los residuos estudentizados demuestra que en líneas generales los residuos siguen una distribución normal al igual que la distribución típica (1,003). No obstante, se destacan dos frecuencias de residuos (-0,67 y 0; 1 y 1,33) alejadas de la línea de distribución normal.

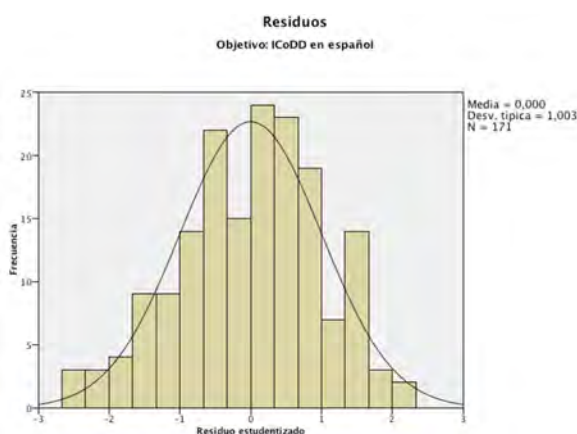


Gráfico 5.38. Histograma de residuos estudentizados del ICoDD en lengua extranjera

La Tabla 5.33 recoge la importancia de cada variable en el modelo de regresión. En primer lugar, se destaca el caudal léxico en lengua extranjera con un coeficiente Beta de 0,303 y una importancia de 0,545. En segundo lugar, se han establecido tres grupos según los resultados de las especialidades: el grupo 1 está

formado por la especialidad de Matemáticas y tiene una incidencia negativa sobre el caudal léxico en lengua española; el grupo 2 lo conforman las especialidades de Lengua Extranjera Inglés y Francés, Lengua y Literatura y Música y tienen una influencia positiva con un nivel de significación de 0,000; el tercer grupo está compuesto por las especialidades de Geografía e Historia, Dibujo, Orientación, FOL, Hostelería y Turismo y TIPI y no ejercen ningún efecto sobre el léxico disponible.

	Coeficiente Beta	Error típico	t	Significación	Nivel de confianza		Importancia
					Inferior	Superior	
Intersección	0,275	0,029	9,608	0,000	0,218	0,332	
ICODD en lengua extranjera	0,303	0,055	5,541	0,000	0,195	0,410	0,545
Especialidad. Grupo 1	-0,050	0,023	-2,135	0,034	-0,134	-0,004	0,347
Especialidad. Grupo 2	0,050	0,014	3,695	0,000	0,023	0,077	0,347
Especialidad. Grupo 3	0,000 ²⁶¹						0,347
ICIT en español	0,006	0,003	1,985	0,049	0,000	0,012	0,070
Hábitos de lectura	0,008	0,006	1,446	0,053	-0,003	0,020	0,037

Tabla 5.33. Datos estadísticos sobre el modelo de regresión lineal del caudal léxico en español

Por orden de importancia le sigue la variable de identificación del léxico especializado con una importancia de 0,070 y los hábitos de lectura con 0,037. Ambas variables tienen una influencia positiva en el caudal léxico y no se distinguen subconjuntos. La variable referida a la lectura presenta un nivel de significación superior al 0,5 razón por la cual en el modelo no se incluye la media estimada.

5.5.2.2.1. Medias estimadas

En el gráfico de medias estimadas del ICoDD en español y el ICoDD en lengua extranjera se evidencia que la relación entre ambas es significativamente positiva. Por consiguiente, se podría decir que aquellos informantes con un mayor caudal léxico en español lo tendrán de forma recíproca en inglés o francés. Asimismo, el ICoDD en lengua extranjera podría servir como variable intermedia ya que transmite la importancia de las especialidades del máster que requieren estudios

²⁶¹ Este coeficiente está establecido en cero porque es redundante.

de lengua extranjera que, a su vez, son las que pertenecen a áreas lingüísticas donde se profundiza en el dominio de un idioma.

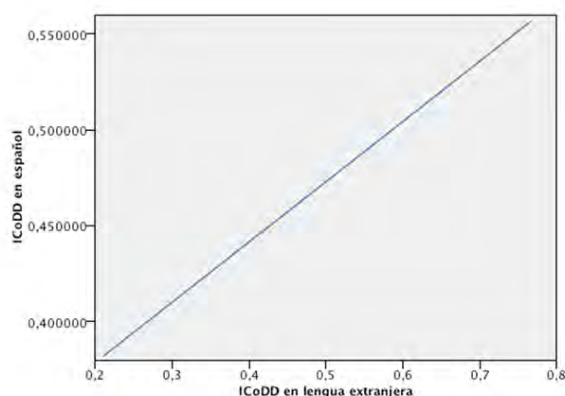


Gráfico 5.39. Medias estimadas del ICoDD en español y el ICoDD en lengua extranjera

a) Especialidad del máster

El Gráfico 5.40 muestra la relación entre la especialidad del máster y el ICoDD en español agrupando los resultados en tres subconjuntos. El primero corresponde a la especialidad de Matemáticas, la cual se sitúa en el mínimo siendo, por tanto, la especialidad que incide negativamente en el ICoDD en español con una media de 0,4119.

A continuación, le sigue el subconjunto que influye positivamente por ser el que presenta mayor ICoDD. Está formado por las especialidades de Lengua Extranjera Inglés y Francés, Lengua Castellana y Literatura y Música y la media estimada es de 0,5107. Por último, se presentan los datos del grupo formado por las especialidades de Geografía e Historia, Dibujo, Orientación, FOL, Hostelería y Turismo y TIPI. La media de este subconjunto es de 0,4645 y no presenta una influencia significativa ni positiva ni negativa.

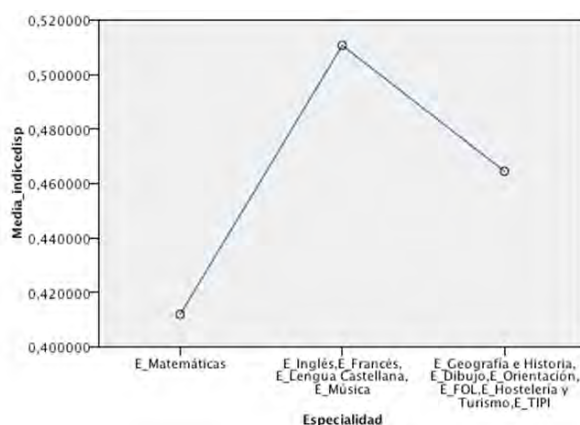


Gráfico 5.40. Medias estimadas del ICoDD en español y la especialidad del máster

b) Léxico especializado en español

El tercer gráfico muestra la relación positiva entre el ICoDD y el ICIT en español. La incidencia de esta variable se justifica por diversas razones. En primer lugar, porque la propia naturaleza del ICoDD tiene en cuenta el vocabulario descentralizado y, por tanto, podría ser especializado. En segundo lugar, porque el hecho de que un informante posea mayor vocabulario quiere decir que también tiene un mejor dominio del lenguaje especializado.

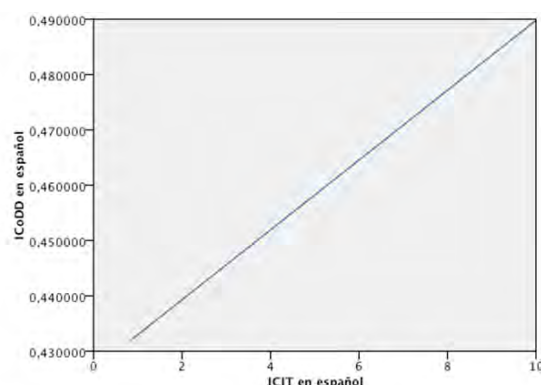


Gráfico 5.41. Medias estimadas del ICoDD y el ICIT en español

Para concluir, cabe destacar que la diferencia de precisión entre el modelo en lengua extranjera y en lengua española es notable. Al buscar una justificación estadística, se constata que la desviación típica y la curtosis es mucho más pequeña en el caso del ICoDD en español que en lengua extranjera lo que muestra que los resultados son más homogéneos. La clase social ocupa un papel secundario que aparece significativo al eliminar valores como el índice de descentralización léxica en español. El resto de variables han resultado ser poco significativas en este análisis o redundantes.

5.5.2.3. Análisis de regresión lineal del léxico especializado en lengua extranjera

A través de la variable del índice corregido de identificación de tecnicismos, se pretende evaluar el conocimiento del vocabulario pasivo en lengua extranjera y en español. Se trata de una prueba complementaria al léxico disponible que evalúa la

capacidad de relacionar palabras especializadas con los nueve centros de interés²⁶² seleccionados para este estudio. Como hipótesis de partida se presupone que uno de los factores que influyen en el vocabulario pasivo es el grado de estudios alcanzado sobre todo en lengua española. Además, en inglés y francés inciden otros factores relacionados con el aprendizaje de la lengua.

Las variables significativas en el modelo de regresión lineal son el índice corregido de identificación de tecnicismos en español, el uso de la lengua extranjera en diferentes contextos, la lengua de la prueba y la clase social. El modelo se explica con una precisión de solo el 43,5 % y un criterio de información de 121,352. Considerando esa escasa exactitud, a continuación se comentan las variables más significativas pero no todas las que podrían explicar el modelo.

El pronóstico por observado dibuja una línea ascendente cuyos datos se concentran entre 0 y 7,5. Aunque no se trata de una recta perfecta de 45 grados, el pronóstico es adecuado. Existen ciertos valores que se alejan de esa línea imaginaria, tales como el ICIT con una puntuación cercana a 10.

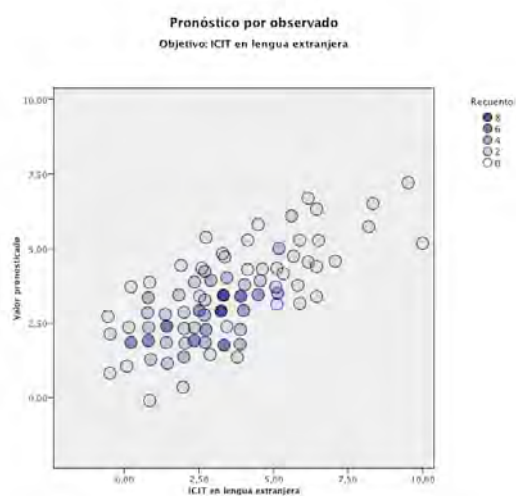


Gráfico 5.42. Pronóstico por observado del ICIT en lengua extranjera

El histograma de residuos estudentizados muestra que las frecuencias de los residuos están bastante próximas a una distribución normal con una desviación típica de 1,003. Sin embargo, el valor de 0 a 0,5 se aleja considerablemente de la media.

²⁶² Vid. apartado 3.1.4 donde se explica en detalle la recogida de datos de la prueba de identificación del léxico especializado.

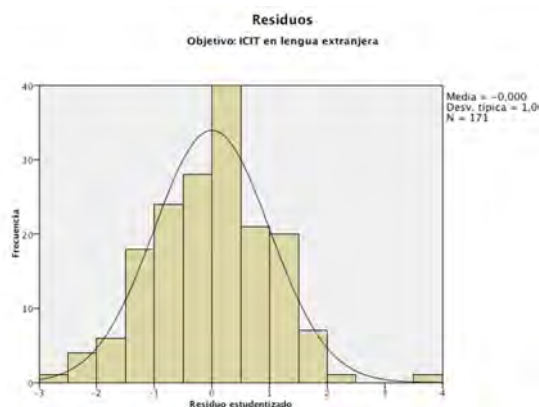


Gráfico 5.43. Histograma de residuos estudentizados del ICIT en lengua extranjera

El coeficiente Beta resultante de la intersección se acerca a uno (0,856) pero el nivel de significación es superior al 0,05 exactamente 0,221, siendo más preciso cuanto mayor se aproxima a cero. El caudal léxico en lengua extranjera es la variable con mayor importancia (0,322) de forma que cuanto más descentralizado es el léxico disponible mayor es el conocimiento del léxico especializado. La siguiente variable por orden de relevancia en el modelo es el uso de la lengua extranjera en diferentes contextos (0,203) que es el resultado del análisis factorial que engloba la realización de estancias en el extranjero, la exposición a los medios de comunicación y el uso de la lengua extranjera en contextos familiares y académicos.

La capacidad de identificación de tecnicismos en español también cobra importancia en este modelo, concretamente 0,203. En cuanto al idioma, el inglés influye negativamente en el caudal especializado ya que presenta un promedio de 2,87 que hace bajar la media global del ICIT que es 3,08. En cambio, la media en francés es superior, 4,54 por lo que produce el efecto contrario. Por último, se sitúa la clase social cuya influencia es menor pero significativa (0,083) y otras variables secundarias que el modelo no ha incluido por considerarla redundantes como es la especialidad del máster y, más concretamente la especialidad del máster en lengua extranjera.

	Coeficiente Beta	Error típico	t	Significación	Nivel de confianza		Importancia
					Inferior	Superior	
Intersección	0,856	0,697	1,229	0,221	-0,519	2,231	
ICoDD en español	4,232	1,082	3,910	0,001	2,095	6,369	0,322
Uso de la lengua extranjera en diferentes contextos	0,465	0,140	3,329	0,001	0,189	0,742	0,234
ICIT en español	0,201	0,065	3,103	0,002	0,073	0,328	0,203
Lengua de la prueba: inglés	-1,009	0,368	-2,741	0,007	-1,736	-0,282	0,158

Lengua de la prueba: francés	0,000 ²⁶³						0,158
Clase social	0,272	0,137	1,981	0,049	0,001	0,543	0,083

Tabla 5.34. Datos estadísticos sobre el modelo de regresión lineal del léxico especializado en lengua extranjera

5.5.2.3.1. Medias estimadas

a) Identificación de tecnicismos en lengua española

La hipótesis inicial, que aquí se corrobora, postula que cuanto mayor dominio del léxico técnico tenga un informante en un idioma mejores serán los resultados de la prueba de identificación de tecnicismos en la otra. El vocabulario científico en las tres lenguas del estudio procede en su mayoría del latín y del griego lo que explica que si el sujeto tiene un dominio en un idioma también influirá en el resto.

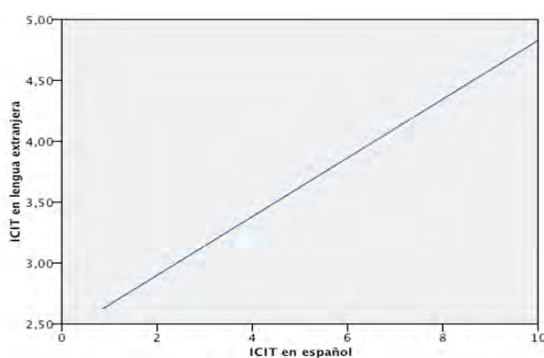


Gráfico 5.44. Medias estimadas del ICIT en lengua extranjera y española

b) Caudal léxico en lengua extranjera

El índice de descentralización léxica mide el vocabulario disponible de los informantes y la actualización de palabras más especializadas que no son compartidas por todos los sujetos. Si los hablantes son capaces de aportar vocablos técnicos también podrán identificar y relacionar unidades léxicas con los centros de interés a los que se refieren. Por tanto, teóricamente el ICIT y el ICoDD son variables relacionadas desde el plano de cualitativo y estadístico. Ya se comentó en el análisis de regresión lineal del ICoDD en lengua extranjera que existe una correlación positiva con el ICIT en lengua extranjera de 0,438 a un nivel de significación bilateral inferior a 0,01.

²⁶³ Este coeficiente está establecido en cero porque es redundante.

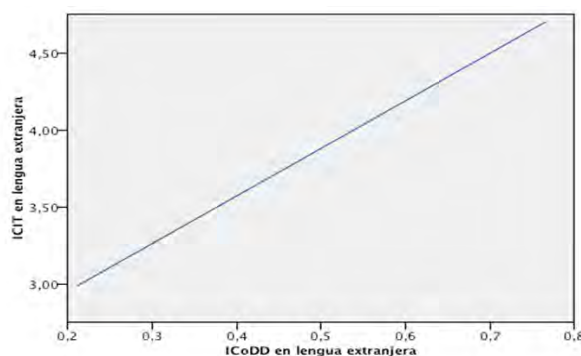


Gráfico 5.45. Medias estimadas del ICIT y el ICoDD en lengua extranjera

c) Uso de la lengua extranjera

El contacto con la lengua extranjera y su uso en situaciones comunicativas diversas son factores claves en el proceso de adquisición de una lengua extranjera. La correlación con la variable de identificación de tecnicismos es significativa ya que cuenta con un coeficiente de correlación de Pearson de 0,494 a un nivel de significancia bilateral inferior a 0,01.

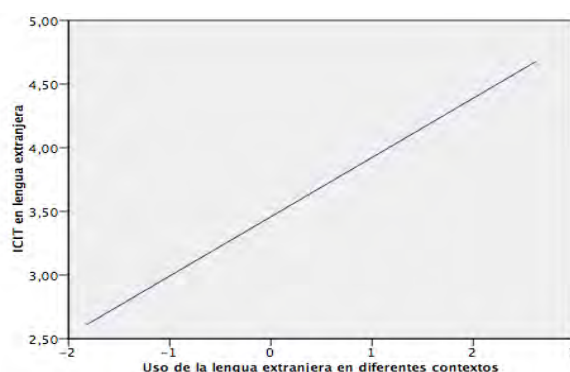


Gráfico 5.46. Medias estimadas del ICIT en lengua extranjera y las asignaturas obligatorias en lengua extranjera en la carrera universitaria

d) Lengua de la prueba

Se produce una diferencia considerable entre los resultados de las pruebas de identificación de tecnicismos en las dos lenguas extranjeras. En francés la media estimada es casi un punto superior al inglés por lo que el coeficiente Beta del idioma germánico es negativo. Aunque el léxico especializado sea compartido en un alto porcentaje en los tres idiomas, el vocabulario común en español se aproxima más al léxico en francés que en inglés. Por esta razón, los sujetos han obtenido puntuaciones más altas en la prueba realizada en francés que en inglés en el léxico pasivo pero no se da tal diferencia en el léxico activo. La correlación bivariada es de 0,294 con un nivel de significancia inferior al 0,01.

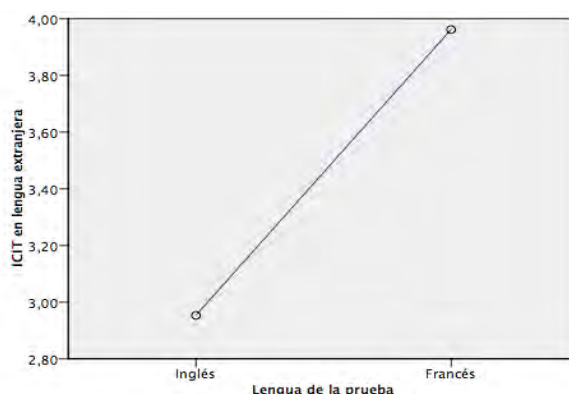


Gráfico 5.47. Medias estimadas del ICIT en lengua extranjera y la lengua de la prueba

e) Clase social

El estatus social juega un factor secundario pero destacado en una muestra donde todos los informantes son postgraduados. La correlación con el índice corregido de identificación de tecnicismos es significativa a un nivel inferior de 0,01 y el coeficiente de Pearson es de 0,367. La clase social influye en factores relacionados con el aprendizaje de un idioma extranjero como es la realización de estancias en el extranjero o la realización de viajes y, por consiguiente, su influencia se traslada al léxico especializado.

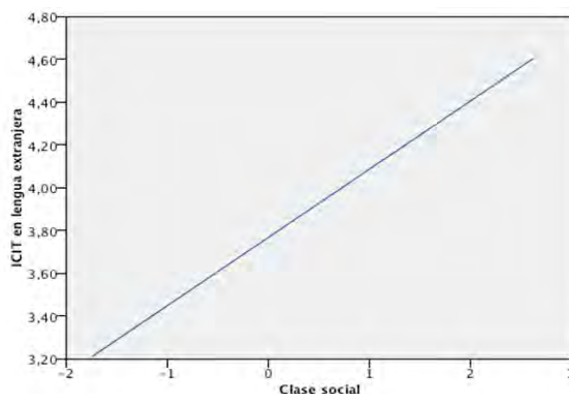


Gráfico 5.48. Medias estimadas del ICIT en lengua extranjera y la clase social

Para concluir, este sería el modelo final de regresión lineal que muestra la influencia de ciertas variables en el ICoDD. Los factores que influyen negativamente son algunos tipos de especialidades (Matemáticas, Orientación, Hostelería, Turismo y TIPI), la variable que y la inseguridad del informante para impartir clases bilingües. En cambio, existen otros factores que inciden de forma positiva, tales como, la especialidad del máster en Lengua Extranjera Inglés, el ICoDD en español y el uso de la lengua extranjera en contexto.

5.5.2.4. Análisis de regresión lineal del léxico especializado en lengua española

El análisis de regresión lineal en lengua española se ha llevado a cabo con numerosas variables y en repetidas ocasiones. No obstante, el nivel de precisión alcanzado ha sido solo del 29,6 % y el criterio de información es de 150,272 por lo que la varianza del modelo no explicaría de forma adecuada la relación entre la capacidad de identificación de tecnicismos en español con respecto al resto de variables. A pesar de ello, se comentan los resultados de este modelo haciendo énfasis en las variables que de forma parcial influyen en el ICIT.

El valor de pronóstico observado presenta una distribución inadecuada ya que el recuento de los valores parece trazar una línea recta. Cabe recordar que un pronóstico óptimo debería tener una inclinación de 45 grados.

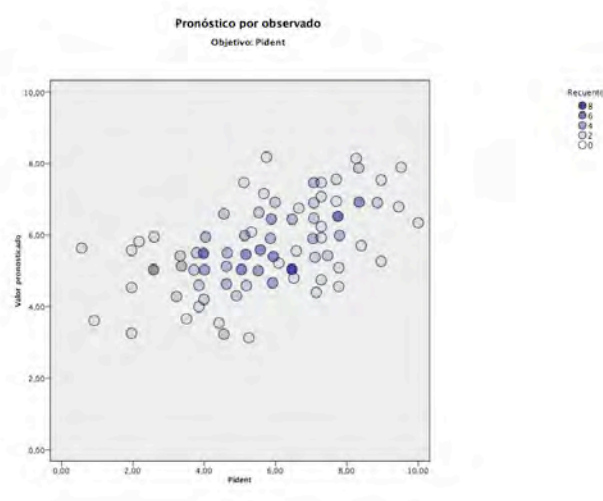


Gráfico 5.49. Pronóstico por observado del ICIT en español

El histograma de residuos estudentizados muestra que, *grosso modo*, la distribución de los residuos coincide con una distribución normal. La desviación típica es adecuada y presenta valores idénticos a la desviación del modelo de regresión de la capacidad de identificación de tecnicismos en lengua extranjera, concretamente de 1,003.

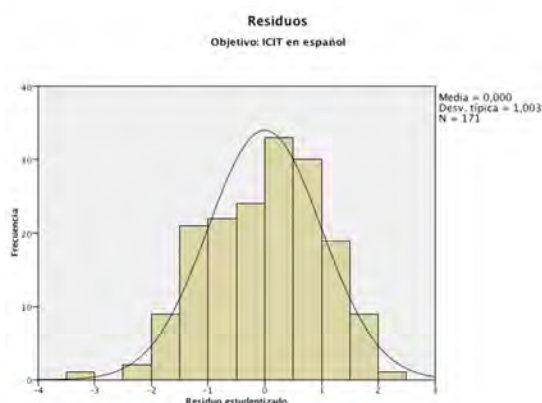


Gráfico 5.50. Histograma de residuos estudentizados del ICIT en español

Los efectos que se destacan en este modelo son la especialidad del máster, el ICIT en lengua extranjera, los hábitos de lectura y la frecuencia con la que los informantes escuchan en la radio programas culturales e informativos. La variable que cobra más fuerza es la especialidad del máster con una importancia de 0,608, seguida por el léxico especializado en lengua extranjera (0,116). A continuación, se sitúan los hábitos de lectura (0,116) y la exposición a programas culturales (0,070) cuya influencia por grupos se analiza más adelante teniendo en cuenta las medias estimadas.

	Coeficiente Beta	Error típico	t	Significación	Nivel de confianza		Importancia
					Inferior	Superior	
Intersección	4,588	0,442	10,369	0,000	3,714	5,461	
Especialidad. Grupo 1²⁶⁴	-1,174	0,514	-2,286	0,024	-2,188	-0,160	0,608
Especialidad. Grupo 2	1,381	0,261	5,303	0,000	0,867	1,896	0,608
Especialidad. Grupo 3	0,000 ²⁶⁵						0,608
ICIT lengua extranjera	0,239	0,067	3,565	0,000	0,107	0,372	0,206
Hábitos de lectura	0,351	0,131	2,678	0,008	0,092	0,610	0,116
Programas culturales. Grupo 1	-0,709	0,340	-2,085	0,039	-1,380	-0,038	0,070
Programas culturales. Grupo 2	0,000 _a						0,070

Tabla 5.35. Datos estadísticos sobre el modelo de regresión lineal del léxico especializado en español

²⁶⁴ Grupo 1: Matemáticas. Grupo 2: Lengua Extranjera Francés; Tecnología, Informática y Procesos Industriales. Grupo 3: Lengua Extranjera Inglés; Lengua y Literatura, Latín y Griego; Geografía e Historia, Dibujo, Imagen y Artes Plásticas; Música; Orientación Educativa; Formación y Orientación Laboral; Turismo.

²⁶⁵ Este coeficiente está establecido en cero porque es redundante.

5.5.2.4.1. Medias estimadas

a) Especialidad del máster

La incidencia de las especialidades en el caudal léxico en lengua española es llamativa ya que presenta una distribución diferente con respecto a los modelos de regresión del caudal léxico a excepción de la especialidad de Matemáticas que presenta el nivel más bajo en ambas pruebas. En cambio, se crean dos subgrupos en las que las especialidades pertenecen a áreas diferentes. El primero está formado por las especialidades de Lengua Extranjera Francés y Tecnología, Informática y Procesos Industriales. De esta forma se asocia, por un lado, una especialidad de área lingüística y, por otro lado, una perteneciente al área científico-tecnológica por contar con un léxico más especializado. El segundo subgrupo está formado por el resto de especialidades y no presenta medias significativamente distintas.

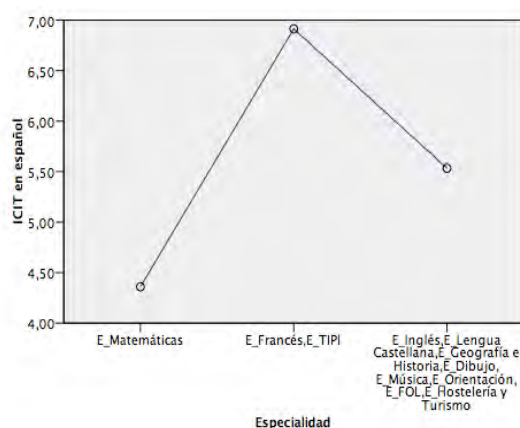


Gráfico 5.51. Medias estimadas del ICIT en español y la especialidad del máster

b) Léxico especializado en lengua extranjera

Al igual que ocurre en lengua extranjera, en este caso se presenta como efecto significativo el ICIT del idioma extranjero. Por tanto, cuanto mayor sea la capacidad de identificación de tecnicismos en lengua española mayor lo será en la lengua extranjera y viceversa.

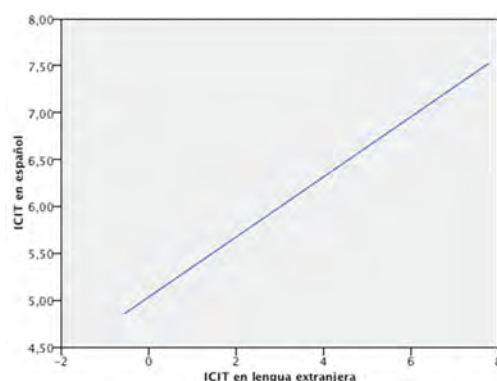


Gráfico 5.52. Medias estimadas del ICIT en español y en lengua extranjera

a) Exposición a los medios de comunicación

Esta variable contabiliza el número de horas que el informante se expone semanalmente a programas de radiofónicos culturales e informativos. Las respuestas siguen el siguiente orden ascendente: ninguna (0), menos de dos (1), entre dos y cinco (2) y más de cinco (3). Se han creado de forma automática dos subgrupos: el primero comprende las tres primeras respuestas y el segundo la última. Cuanto mayor es el número de horas de exposición a los programas culturales mejor será el resultado en el ICIT.

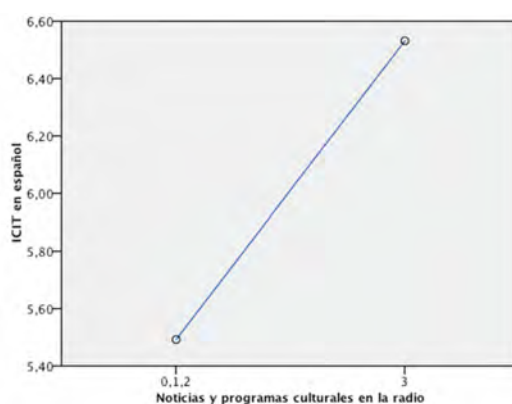


Gráfico 5.53. Medias estimadas del ICIT en español y las noticias y programas culturales en la radio

b) Hábitos de lectura

El siguiente factor que influye positivamente es el número de libros que el informante lee en español al año. El hablante marcó una de las cinco opciones dadas, cuya codificación es la siguiente: ninguno (0), alguno (1), más de cinco (2), más de diez (3) y más de veinte (4). En esta ocasión los datos no se han reagrupado sino que se establece una línea ascendente que relaciona el ICIT con el número de libros.

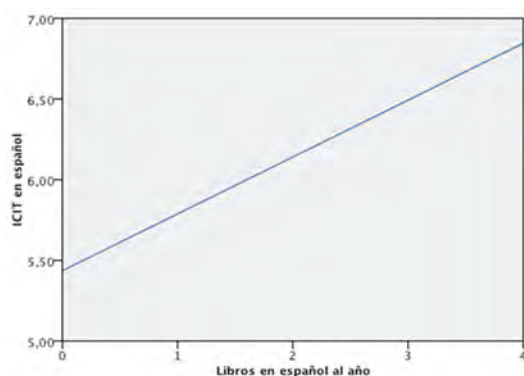


Gráfico 5.54. Medias estimadas del ICoDD en español y la especialidad del máster

En síntesis, el tercer y cuarto análisis multivariantes presentan un nivel de precisión considerablemente inferior al mostrado en el apartado de la variable lingüística. En lengua extranjera el modelo se explica en un 44,2 % y en español apenas alcanza un 14,6 %. A la vista de los resultados, se han comentado los factores que inciden significativamente en las variables ICIT'.

En líneas generales el ICIT' en lengua extranjera está correlacionado de forma positiva con el ICIT' en español, el ICoDD en lengua extranjera, el bilingüismo, las asignaturas cursadas en lengua extranjera durante los estudios universitarios, el uso de la lengua extranjera con amigos y la motivación del informante para trasladarse al extranjero. En cambio, el ICIT' en lengua española está relacionado con menos factores, concretamente con el ICIT' en lengua extranjera, con la asiduidad con la que el informante escucha noticias y programas culturales en la radio y lee libros.

5.5.3. Análisis causales

A través de los análisis de regresión multivariantes anteriores, se han analizados los factores que inciden directamente en las variables dependientes del estudio referidas al caudal léxico y la capacidad de identificación de tecnicismos en español y en lengua extranjera (ICoDD e ICIT'). El análisis causal o análisis de camino es una técnica que complementa los análisis de correlaciones y regresiones parciales y múltiples (García Ferrando, 1989: 413). Para explicar el modelo teórico, el análisis causal se sirve de ecuaciones de regresión múltiple estandarizadas, razón por la cual el diseño de los análisis de este apartado se ha apoyado en los resultados precedentes. El objetivo de los análisis ulteriores es doble:

- Por un lado, determinar cuantitativamente cuál de las variables exógenas tiene mayor influencia en las variables endógenas: índice corregido de descentralización léxica (ICoDD) e índice corregido de identificación de tecnicismos (ICIT) en español y en lengua extranjera.
- Por otro lado, valorar en qué medida las variables sociodemográficas, reticulares y relacionadas con el aprendizaje de una lengua explican el comportamiento de las variables léxicas.

Lévy Mangin y Oubiña Barbolla (2006: 191) clasifican las variables de un modelo causal en variables exógenas y variables endógenas. Dentro de este último grupo distinguen entre las variables dependientes, aquellas que reciben efectos de otras variables, y las variables intermedias, las que muestran relaciones entrantes y salientes. Los investigadores anteriores definen las variables intermedias como «variables situadas dentro de un orden causal entre la variable dependiente y el factor de estudio, pero a diferencia de los factores de confusión forman parte de la cadena causal» (Lévy Mangin y Varela Mallou, 2006: 191). Los datos extraídos de cada modelo son los que siguen:

- Medidas de adecuación del modelo tales como el coeficiente de correlación, el chi-cuadrado y el índice del radical del error de aproximación.
- Los efectos directos que muestran la influencia de una variable sobre otra a través del coeficiente de camino que relaciona ambas variables. En el diagrama se representan con la letra p seguida del valor en cada caso.
- Los efectos indirectos calculados a través del producto de los coeficientes de camino que marcan el trayecto (no directo) de una variable a otra (García Ferrando, 1989: 423).
- Los efectos totales son el compendio de los efectos directos e indirectos. Cabe precisar que los tres tipos de efectos se expresan en forma de coeficientes tipificados Beta.
- El error estándar simbolizado por la letra e .

Con objeto de alcanzar los objetivos perseguidos, se han diseñado dos modelos que corresponden al análisis del ICoDD en lengua extranjera y en español.

La construcción de este modelo causal trata de proporcionar una respuesta a los interrogantes iniciales relacionados con el ámbito sociolingüístico: determinar en qué medida influyen las variables sociales, reticulares y relacionadas con el aprendizaje de una lengua en el léxico disponible en español y en lengua extranjera.

5.5.3.1. Modelo causal de los efectos del hablante sobre el caudal léxico en lengua extranjera

La construcción de este modelo se ha realizado en base a las hipótesis de estudio iniciales sobre la existencia de factores que inciden en el caudal léxico en lengua extranjera de forma directa e indirecta²⁶⁶. En los análisis bivariantes se analizaron las variables con mayor influencia en el caudal léxico y se comprobó la existencia de la relación lineal. Posteriormente, en el modelo de regresión lineal se seleccionaron las variables con mayor incidencia en la capacidad léxica en lengua extranjera: la especialidad del máster, el caudal léxico en lengua española, la motivación para impartir clases bilingües y el uso de la lengua extranjera en diferentes contextos.

La especialidad del máster se distribuye en el modelo de regresión en cinco grupos según el diferente comportamiento sobre el ICoDD²⁶⁷. Los conjuntos formados por la especialidad de Lengua Extranjera Inglés (1) y por las especialidades de Lengua Extranjera Francés y Lengua y Literatura, Latín y Griego (2) influyen positivamente en el caudal léxico ya que cuentan con un media del ICoDD de 0,625 y 0,534 respectivamente. Las especialidades de Geografía e Historia, Dibujo, Imagen y Artes Plásticas, Música y Formación y Orientación Laboral (3) se mantienen en torno a la media, con un índice de 0,468. Por último, la especialidad de Matemáticas (4) junto con el grupo formado por Orientación Educativa, Hostelería y Turismo y Tecnología, Informática y Procesos Industriales (5) inciden de forma negativa en el ICoDD con un promedio de 0,389 y 0,323. La nueva categorización de estas variables se ordena de mayor a menor índice corregido de descentralización léxica.

²⁶⁶ Todas las variables relacionadas guardan una relación lineal que ha sido analizada en el apartado 5.4, referido a los análisis bivariantes y en el apartado 5.5.3.1, donde se expone el modelo de regresión lineal.

²⁶⁷ Se realizaron las pruebas *post hoc* de Tukey y Scheffé para comparar las medias de los grupos y analizar la distribución de las once variables en los cinco subconjuntos homogéneos. *Vid.* apartado 5.5.2.3.1

El caudal léxico en español forma parte de un continuo que recoge los resultados de la variable del índice corregido de descentralización léxica. La motivación para impartir clases bilingües es el resultado de un análisis factorial que tiene en cuenta tres variables: la sensación al impartir clases en lengua extranjera, la opinión sobre la formación en ese aspecto y la predisposición para trabajar en un centro bilingüe. Del mismo modo, el uso de la lengua en diferentes contextos es la variable resultante de un análisis de componentes principales compuesto por otras variables que recogen la asiduidad con la que utilizan los informantes la lengua extranjera en diferentes contextos, la exposición a los medios de comunicación y la realización de estancias en el extranjero.

Una vez incorporadas las variables al modelo de relaciones causales, se ha procedido a evaluar la idoneidad del modelo. El coeficiente de correlación, que oscila entre 0 y 1, es 0,656 por lo que explica con un alto grado de predicción el caudal léxico en lengua extranjera a partir de las variables independientes. En cuanto al chi-cuadrado, es sensible al tamaño muestral pero en líneas generales para indicar buen ajuste debe ser no significativo y contar con valores inferiores a 2 y en este caso es de 1,120. En esta misma línea, se ha verificado la adecuación del modelo a través de otros criterios, tales como el índice del radical del error de aproximación medio (RMSEA: *Root Mean Square Error Aproximation*), el criterio de información de Akaike (AIC: *Akaike Information Criterion*), el índice residual estandarizado (SRMR: *Standardized Root Mean Residual*), entre otros.

A través del análisis de camino es posible explicar los efectos de un conjunto de variables sobre la causa o variable dependiente mediante tres tipos de vías: el camino entre la variable independiente a la intermedia, el camino de la variable intermedia a la dependiente y el resto de caminos relacionados con la variable dependiente o final. En primer lugar, se exponen los efectos sobre las dos variables intermedias en este modelo: la especialidad del máster y el caudal léxico en español. Ambas son clave para la construcción del modelo ya que están fuertemente correlacionadas con la variable dependiente y actúan como variable de mediación con respecto al caudal léxico en lengua extranjera a través de terceras variables.

La motivación tiene un efecto directo sobre la especialidad del máster ya que se presupone que aquellos sujetos que muestran una mayor motivación hacia el bilingüismo eligen estudios universitarios relacionados con áreas lingüísticas donde poder desarrollar sus pretensiones laborales. El efecto directo de la motivación sobre la especialidad del máster es de -0,066 y el efecto indirecto -0,132. Ambos valores son negativos porque a medida que aumenta la motivación descende la media del ICoDD de los cinco grupos de especialidades.

Sin duda, la variable que guarda mayor relación con la especialidad del máster es el uso de la lengua en diferentes contextos de forma que los subgrupos de especialidades relacionadas con el aprendizaje de una lengua (1 y 2) muestran un mayor uso de la lengua en diferentes contextos. Se podría decir que el interés por comunicarse en otra lengua incide también en la selección de estudios determinados que llevan al futuro profesorado a formarse en este aspecto.

En cuanto a los efectos sobre el caudal léxico en español, la motivación incide de forma directa y positiva ya que como se ha visto los sujetos con más interés en impartir clases bilingües suelen estudiar carreras relacionadas con el aprendizaje de una lengua y dicha formación influye también en el caudal léxico en español. En cuanto al uso de la lengua extranjera, tiene un efecto indirecto en el caudal léxico pero no influye directamente.

	Especialidad del máster			ICoDD en español		
	Efectos directos	Efectos indirectos	Efectos totales	Efectos directos	Efectos indirectos	Efectos totales
Motivación	-0,066	-0,132	-0,198	0,059	-0,132	0,163
Uso de la lengua	-0,518	0,000	-0,518	0,000	0,273	0,273
Especialidad				-0,575	0,000	-0,575

Tabla 5.36. Efectos sobre las variables intermedias en el modelo causal del ICoDD en lengua extranjera

Entre los caminos hacia la variable dependiente, destaca la especialidad con un efecto total de -0,659. Se confirma una vez más la importancia de los estudios previos sobre el caudal léxico en lengua extranjera por encima de cualquier otra variable. A continuación, le sigue la variable del caudal léxico en español con un efecto total y directo igual, de 0,335.

Las variables relacionadas con el aprendizaje de una lengua extranjera tienen una importancia muy destacada en el modelo. La motivación muestra un efecto directo de 0,165 y el uso de la lengua 0,125. Los efectos directos se ven completados

por efectos indirectos que hacen que el efecto total de la motivación sea de 0,322 y el uso de la lengua 0,451.

	Efectos directos	Efectos indirectos	Efectos totales
Especialidad	-0,448	-0,180	-0,659
ICoDD español	0,335	0	0,335
Motivación	0,165	0,157	0,322
Uso de la lengua	0,125	0,326	0,451

Tabla 5.37. Efectos en el modelo causal sobre la variable dependiente ICoDD en lengua extranjera

En el diagrama del proceso causal se observa que las variables referidas al uso de la lengua y la motivación son las variables exógenas que influyen tanto en la variable dependiente, el ICoDD en lengua extranjera, como en las variables independientes, el ICoDD en español y la especialidad en lengua extranjera. Los efectos directos de una variable sobre otra se muestran mediante coeficientes tipificados Beta que van precedidos por la letra p y el error o cantidad de variación inexplicada por el modelo causal por la letra e .

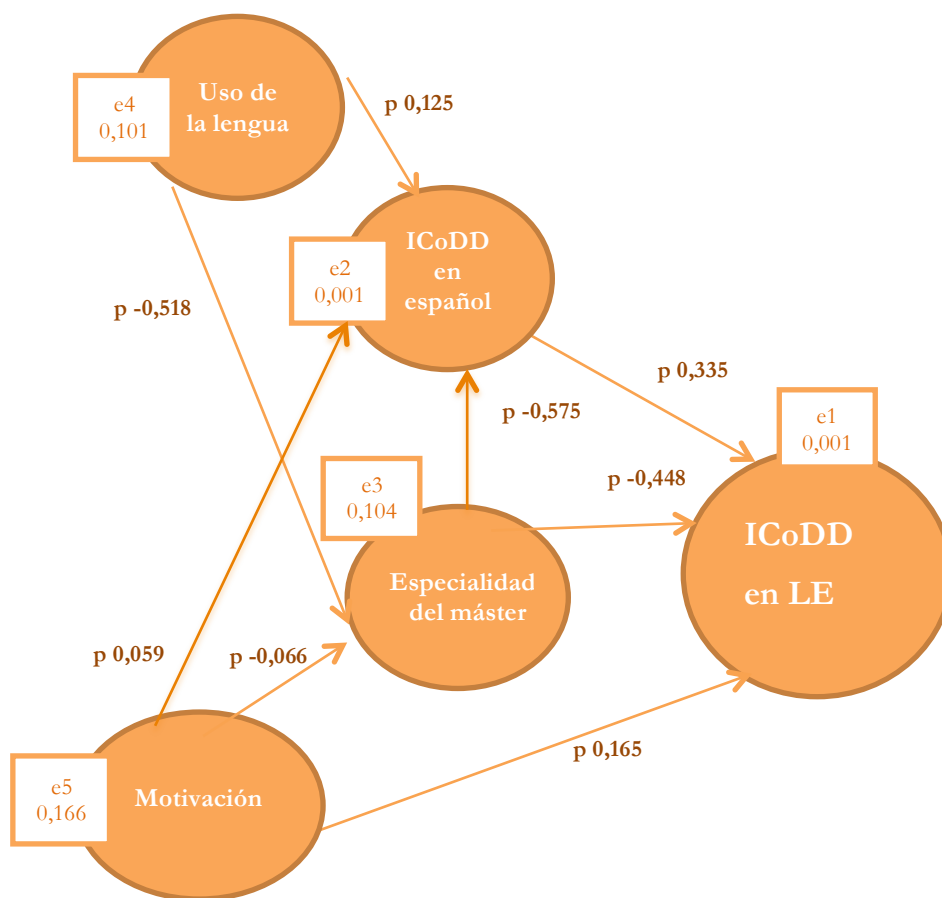


Gráfico 5.55. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICoDD en lengua extranjera

La especialidad del máster es la variable con mayor efecto directo e indirecto sobre el caudal léxico. Se parte del supuesto que la especialidad indica el tipo de estudios previos que ha realizado el informante y, por tanto, son los responsables de que los informantes adquieran más o menos caudal léxico en otra lengua. Sobre la especialidad del máster y sobre el propio caudal léxico influyen directamente la motivación y el uso de la lengua extranjera. El interés o necesidad por usar la lengua en diferentes contextos lleva no solo a los informantes a seleccionar una determinada especialidad sino también a aumentar su caudal léxico en esa lengua.

Por último, el caudal léxico en lengua española recoge la influencia de la motivación para usar una lengua y la motivación por la enseñanza bilingüe porque suele estar muy relacionado con informantes que han realizado estudios de áreas lingüísticas donde es obligatorio el aprendizaje de una lengua extranjera. Cabe precisar que no se han incluido el efecto de variables secundarias como el conocimiento de otras lenguas que se considera redundante al añadir la especialidad del máster y la clase social cuya influencia abarca no solo el caudal léxico sino el hecho de que los informantes hayan realizado estancias en el extranjero o hayan comenzado más prematuramente el aprendizaje de la primera lengua extranjera.

5.5.3.2. Modelo causal de los efectos del hablante sobre el caudal léxico en lengua española

Al igual que en el apartado anterior, se ha construido un modelo de camino cuya variable dependiente es el índice corregido de descentralización léxica en lengua española con el objetivo de conocer en qué medida influyen determinados factores sobre los resultados de la prueba del léxico disponible. Además, este tipo de análisis permite conocer la correlación entre cada par de variables relacionadas ya que en algunos casos al valor directo es necesario sumarle determinados efectos indirectos.

A modo de exordio, se recuerda que el resultado de regresión lineal explica el modelo con un nivel de precisión del 46,3 %. Las variables con mayor incidencia son el ICoDD en lengua extranjera, la especialidad del máster, el índice corregido de identificación de tecnicismos y los hábitos de lectura. En esta ocasión, el modelo divide las especialidades en tres subconjuntos: el primero compuesto por la

especialidad de Matemáticas que hace bajar la media del ICoDD; el segundo formado por las especialidades de Inglés, Francés, Lengua y Literatura y Música que aumenta la media; el tercero y último, formado por las especialidades de Geografía e Historia, Dibujo, Imagen y Artes Plásticas, Orientación Educativa, Formación y Orientación Laboral, Hostelería y Turismo y Tecnología, Informática y Procesos Industriales que se mantiene en torno al promedio del ICoDD.

Los hábitos de lectura son medidos según cuatro intervalos: informantes que no leen ningún libro al año (0); algún libro (1); más de cinco libros (2); y más de diez libros (10). El análisis de regresión muestra una influencia positiva de esta variable sobre el caudal léxico de forma que aquellos informantes que dedican más tiempo a la lectura presentan un léxico activo mayor. En este modelo el efecto directo es de 0,084 y el efecto total es de 0,104. Asimismo, esta variable tiene un efecto directo sobre el índice corregido de identificación de tecnicismos de 0,166 que actúa como variable intermedia.

	ICoDD en lengua extranjera			ICIT en español		
	Efectos directos	Efectos indirectos	Efectos totales	Efectos directos	Efectos indirectos	Efectos totales
Especialidad	-0,640	0,000	-0,640	-0,063	0,000	-0,063
Hábitos de lectura	0,000	0,000	0,000	0,166	0,000	0,166

Tabla 5.38. Efectos sobre las variables intermedias en el modelo causal del ICoDD en lengua extranjera

El ICoDD en español se ve influenciado en un 0,412 por ese mismo índice en lengua extranjera de forma que cuanto mayor es el léxico en una lengua mayor lo será en otra. A continuación le sigue la influencia de la especialidad con un valor negativo, -0,326, ya que los tres grupos de especialidades están organizados de mayor a menor ICoDD. En cuanto al índice corregido de identificación de tecnicismos ejerce una influencia menor, concretamente de 0,119.

	Efectos directos	Efectos indirectos	Efectos totales
Especialidad	-0,326	-0,271	-0,597
ICoDD en lengua extranjera	0,412	0,000	0,412
ICIT	0,119	0,000	0,119
Hábitos de lectura	0,084	0,020	0,104

Tabla 5.39. Efectos en el modelo causal sobre la variable dependiente ICoDD en lengua española

En el diagrama del proceso causal se muestra de forma clara el objetivo, es decir, el ICoDD situado a la derecha y con un error de 0,000. A continuación se sitúan las variables intermedias que son el caudal léxico en lengua extranjera y el

vocabulario especializado en español. Ambos indicadores léxicos han sido analizados en profundidad y se han observado que existen muchas similitudes en el comportamiento de estas variables con respecto a otras extralingüísticas y entre ellas mismas ya que mantienen una relación lineal significativa. Por último, se sitúan las variables exógenas: la especialidad del máster y los hábitos de lectura.

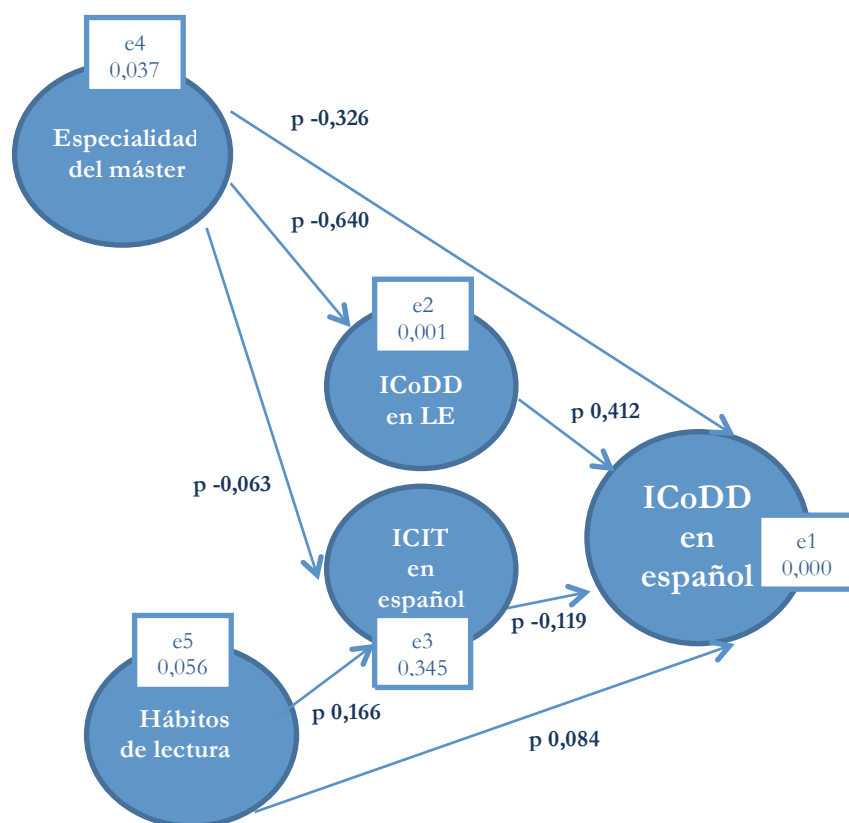


Gráfico 5.56. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICoDD en español

En comparación con los resultados en lengua extranjera, los efectos de las variables sociodemográficas, reticulares y relacionadas con el uso de la lengua se explican con un nivel menor de precisión en las variables lingüísticas en español. Se podría decir que esto se debe a que todos los informantes poseen estudios universitarios y, por tanto, se presupone que ese factor influye de forma significativa en la formación del informante dejando un menor valor explicativo al resto de variables. Asimismo, se ha observado que la variable relacionada con el número de lenguas extranjeras que conoce el informante influye positivamente en el caudal léxico en español pero al añadir el tipo de especialidad del máster resulta redundante

por lo que a pesar de mantener una relación lineal significativa no aparece en el análisis de regresión lineal ni en este análisis causal.

5.5.3.3. Modelo causal de los efectos del hablante sobre el léxico especializado en lengua extranjera

En el análisis de regresión lineal los efectos detectados sobre el índice corregido de identificación de tecnicismos en lengua extranjera son: el ICIT en español, el ICoDD en lengua extranjera, el uso de la lengua, la lengua de la prueba y la clase social. En este modelo se han simplificado algunas variables y se han tomado como referencia las siguientes:

- Variables endógenas: la variable dependiente es el ICIT en lengua extranjera y las variables intermedias son el ICIT en español y el uso de la lengua en diferentes contextos.
- Variables exógenas: la clase social y la especialidad del máster en lengua extranjera en cualquiera de las dos lenguas extranjeras. Esa especialidad del máster conlleva de forma intrínseca la variable relativa a la obligatoriedad de asignaturas en lengua extranjera y, por tanto, la mayor exposición a textos en inglés o francés.

El coeficiente de correlación del ICIT en lengua extranjera de 0,453 y un chi-cuadrado de 17,871 por lo que la adecuación del modelo es menor con respecto al modelo de causal del ICoDD. Los efectos totales sitúan la lengua de la prueba (0,238) como el factor más relevante con un valor muy similar a la clase social (0,234). Le sigue el uso de la lengua (0,209), la especialidad del máster (0,200) y el ICIT en español (0,111). La distribución de los efectos directos es similar ya que los efectos indirectos son menores.

En líneas generales se observa que la clase social junto con la lengua de la prueba son los factores con más influencia en la prueba de identificación de tecnicismos. En cambio, en el léxico disponible la variable con mayor peso era la motivación por lo que se observa una diferencia entre lo que supone la producción de un léxico común y la identificación de tecnicismos.

	Uso de la lengua			ICIT en lengua española		
	Efectos directos	Efectos indirectos	Efectos totales	Efectos directos	Efectos indirectos	Efectos totales
Especialidad en LE	0,645	0,000	0,645			
Clase social	0,280	0,000	0,280	0,169	0,000	0,169
Lengua de la prueba				0,169	0,000	0,132

Tabla 5.40. Efectos sobre las variables intermedias en el modelo causal del ICIT en lengua extranjera

	Efectos directos	Efectos indirectos	Efectos totales
Especialidad en LE	0,200	0,151	0,351
Lengua de la prueba	0,238	0,101	0,339
ICIT en lengua española	0,111	0,028	0,139
Uso de la lengua	0,209	0,000	0,209
Clase social	0,234	0,000	0,234

Tabla 5.41. Efectos en el modelo causal sobre la variable dependiente ICIT en lengua extranjera

Por último, se puede concluir que los informantes que han obtenido un mejor ICoDD en lengua extranjera son capaces de identificar palabras especializadas en mayor medida. De esta forma, la prueba del léxico disponible sirve no solo para evaluar el conocimiento del vocabulario activo de un informante sino también para conocer en qué medida es capaz de identificar léxico especializado.

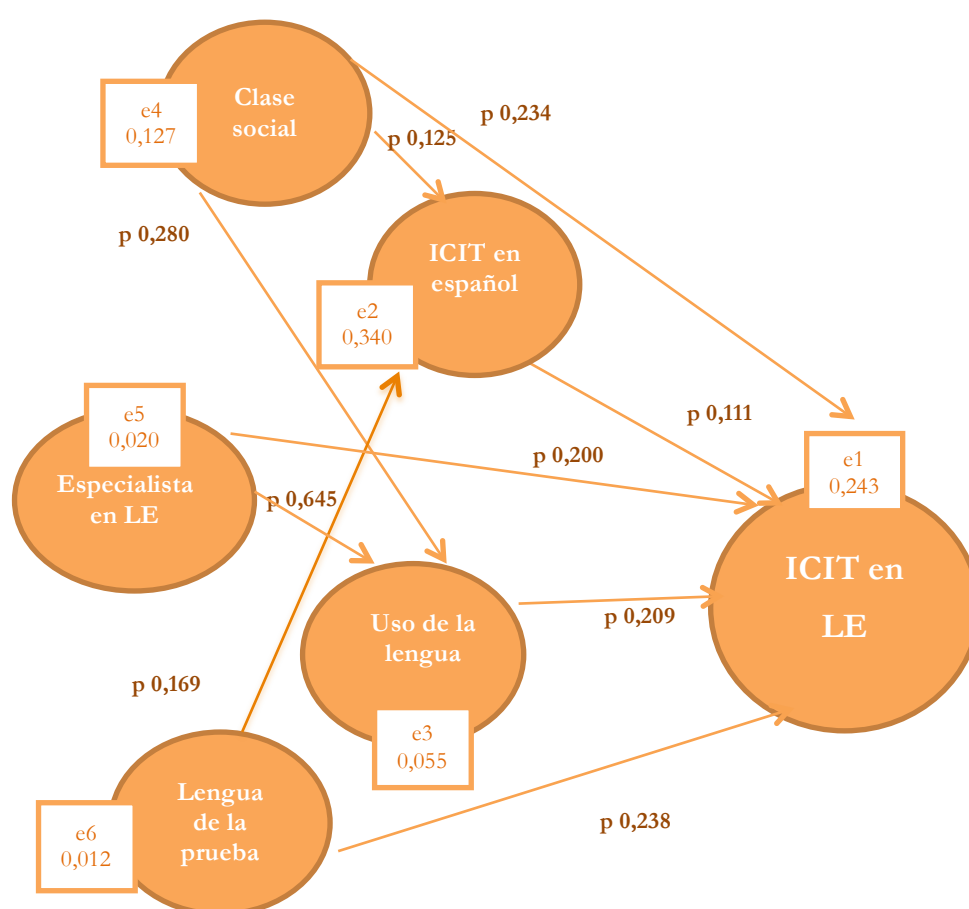


Gráfico 5.57. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICIT en lengua extranjera

5.5.3.4. Modelo causal de los efectos del hablante sobre el léxico especializado en lengua española

El modelo causal sobre el ICIT en español muestra una gran diferencia con respecto a la relación lineal con respecto al coeficiente de determinación. El modelo de regresión lineal apenas se explicaba en un 14,6 % por lo que se limita a presentar aquellas variables que tienen cierta influencia en este índice. Estos factores son el ICIT en lengua extranjera, la frecuencia con la que el informante escucha programas informativos y culturales en la radio, el número de libros que lee al año y la especialidad del máster.

Las variables consideradas en el modelo causal coinciden con el anterior con la salvedad de la inclusión de la clase social cuyo efecto se hace patente sobre el ICIT de forma directa (0,120) e indirecta (0,079). La especialidad se ha dividido en los tres subconjuntos homogéneos localizados en el modelo multicausal ordenados de menor a mayor índice corregido de identificación de tecnicismos. El primer grupo está formado por la especialidad de Matemáticas; el segundo por las especialidades de Lengua Extranjera Francés; Tecnología, Informática y Procesos Industriales; y, el tercero, por el resto de especialidades.

En líneas generales, los índices que muestran la adecuación del modelo causal son débiles. El coeficiente de correlación múltiple es de 0,116 y el chi-cuadrado no es significativo pero es considerablemente superior a niveles considerados adecuados (entorno a 2) concretamente 51,402. El Root Mean Square Error of Approximation RMSEA es bastante elevado 0,264 mientras que un buen modelo presentaría un nivel inferior al 0,05.

Las variables intermedias son el caudal léxico en español, los hábitos de lectura y los programas culturales. La clase social tiene una influencia directa sobre estas tres variables: 0,216 sobre el caudal léxico en español; 0,218 sobre los hábitos de lectura; y 0,175 sobre los programas culturales. De esta forma, la clase social no solo condiciona las variables lingüísticas sino también otros factores extralingüísticos que influyen en la capacidad de identificación de tecnicismos como es la exposición a programas culturales y los libros que lee el informante al año.

La especialidad se ha asociado con el caudal léxico y muestra una influencia positiva de 0,218. En cambio, con respecto a los hábitos de lectura influye negativamente. Cabe destacar que la estructuración de la especialidad se ha realizado en base a los subconjuntos homogéneos creados a partir de la variable de identificación de tecnicismos en español. En este caso la influencia es poco significativa -0,063 y quizás se deba a que los conjuntos de especialidades se reagrupan de forma diversa según los hábitos de lectura. A su vez, esta última variable influye ligeramente menos en la capacidad léxica que en el léxico especializado (0,080 frente a 0,109) mostrando así una mayor influencia en el denominado léxico pasivo con respecto al léxico activo.

	ICoDD en español			Hábitos de lectura			Programas culturales		
	Efectos directos	Efectos indirectos	Efectos totales	Efectos directos	Efectos indirectos	Efectos totales	Efectos directos	Efectos indirectos	Efectos totales
Especialidad	0,218	0,000	0,218	-0,063	0,000	0,126			
Hábitos de lectura	0,080	0,000	0,080						
Programas culturales	-0,042	0,003	-0,038	0,041	0,000	0,041			
Clase social	0,216	0,003	0,218	0,118	0,007	0,126	0,175	0,000	0,175

Tabla 5.42. Efectos sobre las variables intermedias en el modelo causal del ICIT en lengua española

Las variable que presenta mayor influencia directa en el ICIT en español es el caudal léxico (0,186) de forma que se repite esta asociación significativa con respecto al modelo causal del ICoDD en lengua española. A continuación, le siguen los programas y noticias culturales (0,140), la clase social (0,120), los hábitos de lectura (0,109) y la especialidad del máster (-0,103).

	Efectos directos	Efectos indirectos	Efectos totales
ICoDD en español	0,186	0,000	0,186
Especialidad	-0,103	0,000	-0,103
Hábitos de lectura	0,109	0,015	0,124
Programas culturales	0,140	0,003	0,137
Clase social	0,120	0,079	0,199

Tabla 5.43. Efectos en el modelo causal sobre la variable dependiente ICIT en lengua española

El

Gráfico 5.56 muestra la influencia de las variables lingüísticas sobre el ICIT donde además de las variables lingüísticas se observa la importancia de los libros al año y de los programas culturales. En comparación con el modelo del ICoDD, se han incluido los programas culturales e informativos como vía de adquisición de vocablos técnicos sumado a la variable de los hábitos de lectura presente en ambos modelos. El error

estándar del índice corregido de identificación de tecnicismos es bastante elevado 0,321 seguido por la clase social 0,125 y los hábitos de lectura 0,086. El resto de variables presentan errores estándares inferiores al 0,05: la especialidad del máster (0,030), los programas culturales (0,021) y el caudal léxico en lengua española (0,001).

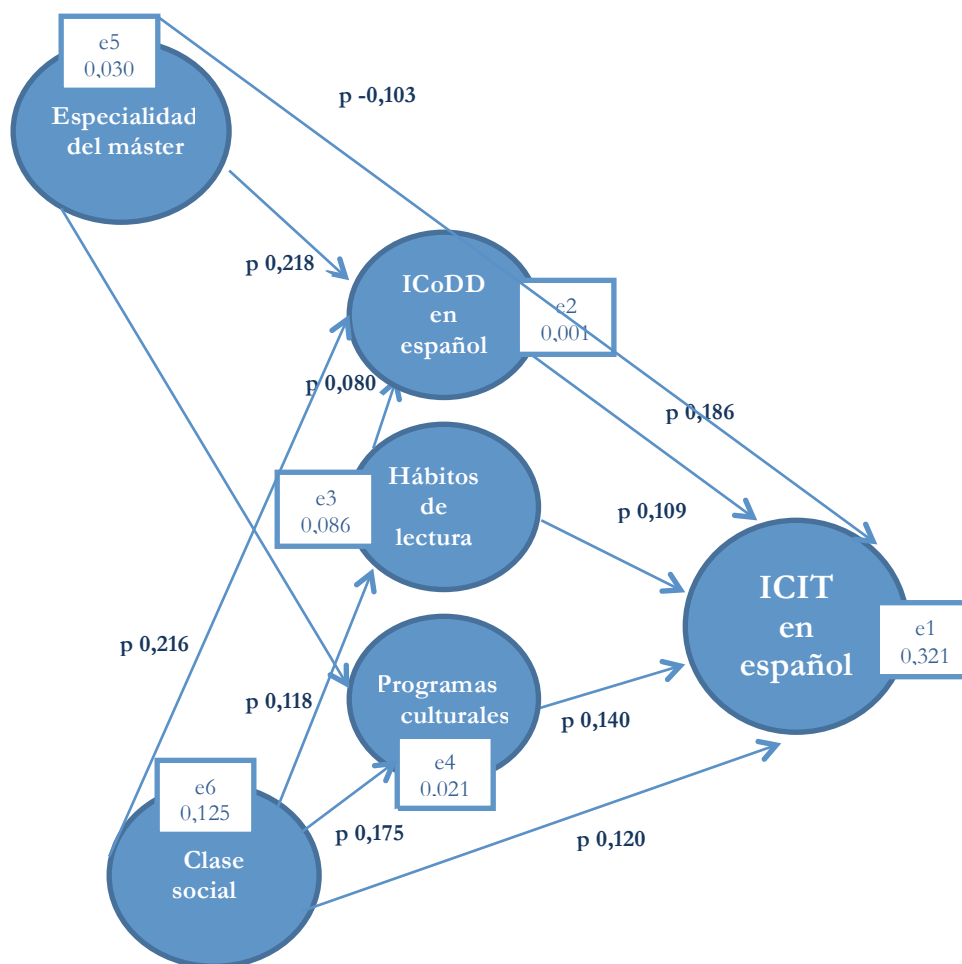


Gráfico 5.58. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICIT en lengua española

Dada los bajos niveles de adecuación de este modelo, los resultados sirven para esbozar qué variables pueden condicionar el léxico especializado: el caudal léxico, los hábitos de lectura, la exposición a programas culturales, la especialidad del máster y la clase social. Se corrobora la hipótesis que muestra que la variable con mayor influencia en el vocabulario pasivo es el nivel de estudios alcanzado por los estudiantes sobre cualquier otra variable de forma que homogeniza los resultados y hace que la incidencia del resto de variables sea inferior.

5.6. Conclusiones parciales

A lo largo de este capítulo, se han presentado y analizado las variables resultantes del cuestionario sociológico, de las pruebas de disponibilidad léxica y de identificación del léxico especializado con un doble objetivo: conocer cuáles son los factores que inciden en el caudal léxico y en la capacidad de identificación de tecnicismos en español y en lengua extranjera. Al mismo tiempo, los resultados pretenden servir de orientación pedagógica tanto para la lingüística aplicada a la enseñanza de idiomas como para la organización universitaria de la formación inicial del futuro profesorado.

En primer lugar, se han presentado las hipótesis del estudio desde la perspectiva de la sociolingüística. Para poder corroborar o refutar los supuestos planteados, se han llevado a cabo una serie de análisis de forma secuenciada y justificada: análisis univariantes para presentar las variables, análisis bivariantes para conocer la relación entre dos variables y análisis multivariantes para conocer la incidencia de las variables extralingüísticas sobre las variables lingüísticas o dependientes: el ICoDD y el ICIT.

El índice corregido de descentralización léxica (ICoDD) mide la capacidad léxica de los informantes a partir de los resultados de la prueba de disponibilidad léxica. Para calcularlo, se han tenido en cuenta las palabras menos compatibles o aportadas por una minoría de informantes y la relación con el centro de interés. De esta forma, se obtiene un índice que da cuenta no solo del número de palabras aportadas por los informantes, sino también de su grado de descentralización y relación con respecto al centro de interés.

En lengua extranjera, los factores que guardan mayor relación con el ICoDD están relacionados en su mayoría con el aprendizaje de una lengua extranjera, además del propio ICoDD en lengua española. En el modelo de regresión lineal y el modelo causal, se han destacado la motivación para impartir clases de enseñanza bilingüe, el uso de la lengua extranjera en su contexto natural y la especialidad del máster que están cursando los informantes. Existen otros factores, como el conocimiento de otras lenguas extranjeras, que guardan una relación significativa con el ICoDD pero

resultan redundantes en los modelos multivariantes por estar relacionados directamente con la especialidad del máster.

En español, tanto los análisis bivariantes como multivariantes, muestran una relación menos significativa entre las variables lingüísticas y extralingüísticas con el ICoDD. Asimismo, el análisis de regresión lineal y el análisis causal explican con un grado de adecuación menor que en lengua extranjera el comportamiento del caudal léxico en español. La principal causa de este efecto se debe al papel principal que juega el grado de formación de los hablantes en la competencia léxica de los hablantes. Como todos los estudiantes son de posgrado, existe un margen menor de incidencia en el caudal léxico que viene marcado por otras variables lingüísticas, el ICoDD en lengua extranjera y el ICIT en español, la especialidad del máster y los hábitos de lectura.

En cuanto al índice corregido de identificación de tecnicismos (ICIT) mide la capacidad del informante de reconocer palabras especializadas y relacionarlas con un centro de interés dado. Se obtiene aplicando una fórmula matemática que tiene en cuenta las respuestas correctas e incorrectas resultantes de la prueba del léxico especializado. El objetivo de esta prueba es conocer el vocabulario pasivo del informante ya que se verifica la capacidad de reconocer ciertas unidades léxicas. En cambio, este hecho no conlleva necesariamente que el sujeto pueda utilizar la palabra en contexto y que forme parte del vocabulario activo.

Los análisis de regresión lineal múltiple y causales, que miden los factores que influyen en la capacidad de identificar palabras especializadas, presentan un modelo de adecuación bastante débil en lengua extranjera y, sobre todo, en español. No obstante, resulta interesante destacar las variables con mayor influencia en el ICIT en lengua materna y extranjera para tenerlas en cuenta en el ámbito de la lingüística aplicada a la enseñanza de lenguas.

En lengua extranjera, las variables con una incidencia positiva sobre el ICIT son el uso de la lengua en diferentes contextos, el ICIT en español, la elección del francés en la prueba, el ICoDD en lengua extranjera y la clase social. Al igual que en los modelos multivariantes del caudal léxico, variables como el conocimiento de otros idiomas guardan una relación significativa bilateral con el ICIT están presentes

en los modelos multivariantes de forma indirecta, como puede ser a través del uso de la lengua y la especialidad del máster.

En lengua española, se produce una relación de influencia recíproca con el ICIT en lengua extranjera. El resto de variables destacadas son los hábitos de lectura, la exposición a programas culturales, la especialidad del máster y la clase social. Por tanto, una de las conclusiones que se pueden extraer desde un enfoque pragmático y didáctico es que el número de libros que un estudiante lee al año influye en su vocabulario activo y pasivo.

CAPÍTULO 6. SELECCIÓN DEL LÉXICO Y PROPUESTA METODOLÓGICA

- 6.1. Introducción**
- 6.2. Organización de la memoria semántica**
- 6.3. La teoría de los prototipos en los estudios de disponibilidad léxica**
- 6.4. Procesos de reentrada del léxico**
 - 6.4.1. Modelos de reentrada según la ontogenia de los centros de interés
 - 6.4.2. Modelo de reentrada de los centros de interés prototípicos
 - 6.4.2.1. Modelo de reentrada leve
 - 6.4.2.2. Modelo de reentrada moderado
 - 6.4.2.3. Modelo de reentradas pronunciado
 - 6.4.2.4. Comparación entre la secuencia léxica de un informante en español y en lengua extranjera
 - 6.4.3. Modelo de reentrada de los centros de interés relacionales
 - 6.4.3.1. Modelo de reentrada moderado
 - 6.4.3.2. Modelo de reentrada pronunciado con desviación lineal
 - 6.4.4. Ejemplo de clasificación del centro de interés *Ordenadores e internet* según los procesos de reentrada
 - 6.4.4.1. Modelo de reentrada moderado
 - 6.4.4.2. Proceso de reentrada pronunciado con desviación lineal
- 6.5. Análisis de las relaciones semánticas a través del DispoGrafo**
 - 6.5.1. Análisis de las relaciones semánticas según el centro de interés
 - 6.5.2. Organización del léxico según el idioma
 - 6.5.2.1. Resultados del centro denominado *Partes del cuerpo humano*
 - 6.5.2.2. Resultados del centro denominado *La escuela: muebles y materiales*
 - 6.5.2.3. Resultados del centro denominado *Ordenadores e internet*
- 6.6. Hacia una propuesta didáctica**
 - 6.6.1. Selección léxica según diferentes criterios de corte
 - 6.6.1.1. Léxico según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición
 - 6.6.1.2. Determinación del léxico relevante
 - 6.6.2. Los aspectos socioculturales en el léxico disponible

6.6.3. El léxico nuclear y el nivel según el Marco Común Europeo de Referencia para las Lenguas

6.6.3.1. Léxico nuclear en los idiomas del estudio

6.6.4. Léxico nuclear en el Plan Curricular del Instituto Cervantes

6.6.5. Propuesta didáctica

6.7. Conclusiones parciales

6.1. Introducción

Las investigaciones relacionadas con la organización de las palabras en la mente de los individuos tratan de descifrar, mediante diferentes tipos de pruebas, cuáles son los procesos cognitivos y psicolingüísticos que se producen al actualizar una palabra. El objetivo de este capítulo es conocer algunos de esos factores cognitivos con objeto de dotar a los profesionales del ámbito de la enseñanza del español como lengua materna o lengua extranjera de instrumentos para seleccionar el léxico según el nivel de sus aprendices.

En primer lugar, se define la fundamentación teórica de los principales modelos de organización de la memoria semántica: teoría de los rasgos, teoría de las redes semánticas, teoría de los prototipos y postulados de significado. A continuación, se desarrollan los dos modelos que según Hernández Muñoz (2006: 67), están más relacionados con la disponibilidad léxica. De esta forma, se muestran las relaciones entre las palabras de dos formas diferentes pero complementarias.

Por un lado, se presenta el léxico prototípico o nuclear generado a partir del aplicación de la teoría de los conjuntos difusos a los resultados del índice de compatibilidad léxica. Se destaca el vocabulario presente en los listados en español, inglés y francés de un centro de interés para comparar el grado de coincidencia y evaluar el criterio de selección léxica aplicado²⁶⁸. Una vez definido el léxico nuclear resultante de la prueba de disponibilidad, se evalúan los procesos de reentrada al núcleo seleccionando un centro prototípico (*Partes del cuerpo humano*) y un centro relacional (*La escuela: muebles materiales*). Se han tomado como referencia las clasificaciones de los centros de interés tradicionales propuesta por Hernández Muñoz (2006), Sánchez-Saus Laserna (2011) y Tomé Cornejo (2011) para realizar una nuevo esbozo. En este estudio se ha incluido un centro nuevo, *Ordenadores e internet*, que no ha sido analizado en estudios anteriores. Para completar la taxonomía se ha llevado a cabo el análisis de ese centro de interés para poder definir empíricamente su estructura.

²⁶⁸ El resto de listados con el léxico compatible de todos los centros de interés en los tres idiomas se encuentra en el Anexo VIII.

Por otro lado, se realiza un análisis de las relaciones mediante el programa DispoGrafo. El objetivo es mostrar de forma visual cuáles las conexiones entre los diferentes nodos o vocablos y el peso o fuerza de esas relaciones en los tres idiomas del estudio. Los resultados permitirán conocer las relaciones de cada unidad léxica y su grado de conectividad con objeto de acercar las propuestas didácticas a la red semántica de los informantes.

Seguidamente, se realiza una selección del léxico según tres parámetros: vocablos que han sido aportados por un único informante; lexías según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición y por el *Fuzzy Expected Value*, utilizado como la herramienta capaz de delimitar el léxico relevante.

Por último, se tienen en cuenta esos resultados para reflexionar sobre la importancia de un programa que permita conocer cuál es el vocabulario más adecuado para los estudiantes según su competencia lingüística. El primer paso reside en conocer la similitud de los centros de interés con respecto a las nociones específicas del Plan Curricular del Instituto Cervantes para evaluar la posible modificación de los centros de interés tradicionales. A continuación, se propone la creación de una herramienta informática que incluya los tres grandes bloques tratados: el léxico nuclear, las relaciones entre las palabras y su clasificación según el Marco Común Europeo de Referencia para las Lenguas.

6.2. Organización de la memoria semántica

Los orígenes de la disponibilidad léxica se remontan a 1951 cuando el Ministerio de Educación Nacional crea una comisión para elaborar un léxico de base con fines didácticos²⁶⁹. Por tanto, se podría decir que el primer objetivo de estos estudios fue crear una lista con las palabras más idóneas para estudiantes de francés. La recogida de datos se realiza a partir de un estímulo determinado o centro de interés sobre el cual los informantes tienen que escribir las palabras que se les vengán a la mente.

²⁶⁹ Vid. apartado 2.2.

En cambio, el procedimiento no está exento de críticas²⁷⁰ desde el punto de vista ideológico (Cohen, 1955) así como metodológico²⁷¹ (Juilland, Brodin y Davidovitch, 1970; Galisson, 1971, 1991). Hernández Muñoz (2006) plantea la cuestión de si realmente se consigue recopilar las palabras más disponibles a través de este tipo de pruebas o si, simplemente, se recogen las palabras que se activan más rápidamente.

Uno de los grandes interrogantes es conocer cómo se organizan las palabras en la mente de los hablantes (Carnap, 1952; Katz y Fodor, 1963; Katz, 1972; Kripke, 1972, Rosch, 1973, 1978; Fodor, Fodor y Garrett, 1975). En este capítulo se abordarán algunos de los factores que ayudan a explicar esa organización léxica con un doble objetivo: conocer a grandes rasgos cómo se han obtenido las diferentes listas de palabras y conseguir seleccionar el vocabulario relevante para la enseñanza de español. Concretamente, se analizarán la teoría de la prototipicidad y las relaciones de las palabras.

Antes de analizar la organización concreta de las palabras del estudio, conviene contextualizar el tema para que, como si de una relación de meronimia se tratase, sean comprendidas como parte de un todo. En primer lugar, se definen de forma breve los conceptos de lexicón mental y memoria semántica.

El lexicón mental, *grosso modo*, puede entenderse como el diccionario o almacén de palabras que recoge tanto las unidades léxicas como su información semántica y sintáctica (Aitchison, 2012). En cambio, según este autor, este concepto puede abordarse desde enfoques más integradores que amplíen los límites de ese almacén a otros ámbitos, tales como el acústico y el fonético (Marslen-Wilson, 1989) o más específicos que limitan su definición a la activación del léxico.

La memoria semántica hace referencia al almacén de conceptos que recoge la información lingüística y extralingüística a nivel cognitivo. Una de las definiciones más difundidas es la aportada por Tulving (1972: 386):

²⁷⁰ Gougenheim (1956) y Rivenc (1956) justifican el proyecto y responden a algunas de las críticas recibidas.

²⁷¹ La mayoría de las reflexiones en este sentido están relacionadas con la idoneidad de los centros de interés (Borrego Nieto, 2004; Hernández Muñoz, 2006; Bartol Hernández 2010; Moreno Fernández, 2012; González Fernández, 2014; Paredes García, 2014; entre otros).

Semantic memory is the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas and algorithms for the manipulation of these symbols, concepts and relations. Semantic memory does not register perceptible properties of inputs, but rather cognitive referents of input signals. The semantic system permits the retrieval of information from the system leaves its contents unchanged, although any act of retrieval constitutes an input into episodic memory. Finally, the semantic system may be quite independent of the episodic system in recording and maintaining information since identical storage consequences may be brought about by a great variety of input signals.

Hernández Muñoz (2006) analiza y describe los factores psicolingüísticos que intervienen en la disponibilidad léxica. Para ello, realiza un minucioso recorrido teórico sobre los principales fundamentos cognitivos y los acompaña de diferentes experimentos psicolingüísticos que ayudan a fundamentar sus conclusiones. La autora dedica un apartado de su tesis doctoral denominado «Organización semántica de la memoria» (2006: 64-68) a recoger los principales modelos de organización de la memoria semántica. La clasificación la realiza a partir de los trabajos de Schreuder y Flores d'Arcais (1989), de Garnham (1992) y distingue estos cuatro tipos²⁷²:

1. Teoría de los rasgos (*feature theory*) o teoría descomposicional. Esta teoría presupone que todo morfema, palabra, frase u oración puede descomponerse en una serie de rasgos o componentes semánticos denominados hormantes semánticos. Este concepto es definido por uno de los precursores de esta teoría como «un constructo teórico destinado a representar un concepto que forma parte del sentido de los morfemas y otros constituyentes de las lenguas naturales» (Katz, 1979: 52).

Los hormantes semánticos forman parte de una lección que se construye a partir de ciertas reglas que determinarán los casos en los que los hormantes se pueden combinar para crear sentidos más complejos. Esta teoría se complementa y se amplía posteriormente con la teoría lingüística de descomposición léxica de Katz y Fodor (1963) y la teoría descomposicional de Schank (1975). En cambio, algunos detractores como Putnam (1975), Kripke

²⁷² Existen otros fundamentos teóricos como por ejemplo la Teoría de la Relevancia (Sperber y Wilson, 1986) pero no se han incorporado en esta clasificación.

(1972), Fodor, Fodor y Garrett (1975), entre otros, critican ciertos aspectos de la teoría.

2. Teoría de las redes semánticas (*semantic networks*). Existen numerosos estudios que consideran que el conocimiento se encuentra organizado en forma de red semántica (Quillian, 1968; Collins y Quillian, 1972; Rumelhart, Lindsay y Norman, 1972). La teoría postula que cada red está formada por un conjunto de nodos o representaciones abstractas unidas mediante conexiones o vínculos inferenciales. La información semántica se activa a partir de las relaciones con otros nodos de tal manera que su accesibilidad dependerá de la fuerza de las conexiones.
3. Teoría de los prototipos (*Prototypes theory*). Esta teoría surge de los estudios realizados por la psicóloga Eleanor Rosch (1973 y 1978) basándose en las reflexiones filosóficas de Wittgenstein (1953). Los elementos semánticos se definen a partir de un núcleo o prototipo que engloba una categoría. Cuanto más cerca se encuentre una unidad léxica del núcleo, mejor será el grado de similitud con que se define la categoría y más accesible se encontrará en la memoria.
4. Postulados de significado (*Meaning postulates*). Esta teoría desarrollada por Fodor, Fodor y Garrett (1975) se basa en los fundamentos teóricos expuestos por Carnap (1952). A diferencia de la teoría de rasgos, las unidades léxicas simples no se pueden descomponer en unidades menores sino que tienen sentido en sí mismas. «Our proposal is that to each morpheme of the surface vocabulary of a natural language, there correspond a primitive expression in the vocabulary of the representational system» (Fodor, Fodor y Garrett, 1975: 525).

Para concluir, Hernández Muñoz (2006: 67) afirma que en el ámbito de la disponibilidad léxica los modelos más presentes son el referido a la teoría de las redes semánticas (López Morales, 1999a; Urrutia Martínez, 2001; Galloso Camacho, 2002a; Gómez Devís, 2004; Ferreira, 2006) y la teoría del prototipo (Romero Rubilar, 2000; Ávila Muñoz y Sánchez Sáez, 2010; López González, 2014) relacionadas con el lexicón mental. No obstante, aboga por el estudio y análisis de los factores que

inciden tanto en el lexicón mental como en la memoria semántica. Razón por la cual en los siguientes apartados se analizará en primer lugar una aplicación práctica de la teoría del prototipo y, en segundo lugar, de la teoría de las redes semánticas.

6.3. La teoría de los prototipos en los estudios de disponibilidad léxica

La teoría de los prototipos (Wittgenstein, 1953; Rosch, 1973, 1978; Lakoff, 1987) ha servido como fundamento teórico para explicar la estructura y la relación de las palabras en la mente en el ámbito de la disponibilidad léxica (Hernández Muñoz, 2006; Sánchez-Sauz Laserna, 2011; Tomé Cornejo, 2011, Ávila Muñoz y Sánchez Sáez, 2010, 2011, 2014; López González, 2014). Ávila Muñoz y Sánchez Sáez (2010, 2011, 2014), parten de la idea de que el léxico aportado por los informantes gira en torno a un núcleo o prototipo. Los vocablos considerados más nucleares o prototípicos se sitúan en el centro de la espiral y son más accesibles a la mayoría de los informantes. A medida que los vocablos se alejan de este núcleo, son menos prototípicos y pierden accesibilidad, razón por la cual suelen presentar un índice de disponibilidad menor.

Anteriormente y con objeto de analizar la relación entre la disponibilidad y la tipicidad, Romero Rubilar (2010) realizó una prueba de disponibilidad léxica y un cuestionario de prototipicidad a un grupo de 40 alumnos de enseñanza media en dos ocasiones con una diferencia de siete meses. En el cuestionario sobre el léxico prototípico, los estudiantes tuvieron que marcar el grado de representatividad (alta, media o baja) de 15 palabras. El cotejo entre las respuestas de los informantes y el grado de disponibilidad léxica de esos vocablos muestra una relación del 84 %. En cambio, los informantes tan solo repiten en la segunda prueba de disponibilidad el 51 % de los vocablos escritos en la primera.

Esa prototipicidad ligada al concepto de disponibilidad puede hacerse visible en los resultados comparativos que se muestran en el capítulo 4. Se observa una gran coincidencia entre los veinte primeros vocablos más disponibles de investigaciones diversas (estudiantes preuniversitarios y postuniversitarios) en lengua española e incluso en diferentes lenguas (léxico en inglés y francés). Ante un estímulo dado, como puede ser *El cuerpo humano*, los informantes de diferentes edades y países

proporcionan palabras muy similares en torno al prototipo solicitado²⁷³. Por tanto, los resultados sirven para confirmar que la estructura interna de ciertos prototipos o centros de interés es universal.

Uno de los grandes retos consiste en definir en qué medida un vocablo es más o menos representativo de un prototipo. Rosch (1978) expone que existen diferentes grados de tipicidad entre los que se sitúan las diferentes unidades léxicas. En este sentido, podría decirse que las palabras disponibles forman parte de un continuo de representatividad con respecto al núcleo que pueden definirse según su compatibilidad léxica. *Grosso modo*, este marcador mide la compatibilidad de cada vocablo y su relación con el conjunto de unidades léxicas aportadas por el todos los informantes que conforman la muestra de estudio. Para calcularlo, Ávila Muñoz y Sánchez Sáez (2010 y 2011) se basan en los preceptos de la teoría de los conjuntos difusos y establecen hasta seis criterios de corte del léxico según su compatibilidad o accesibilidad²⁷⁴.

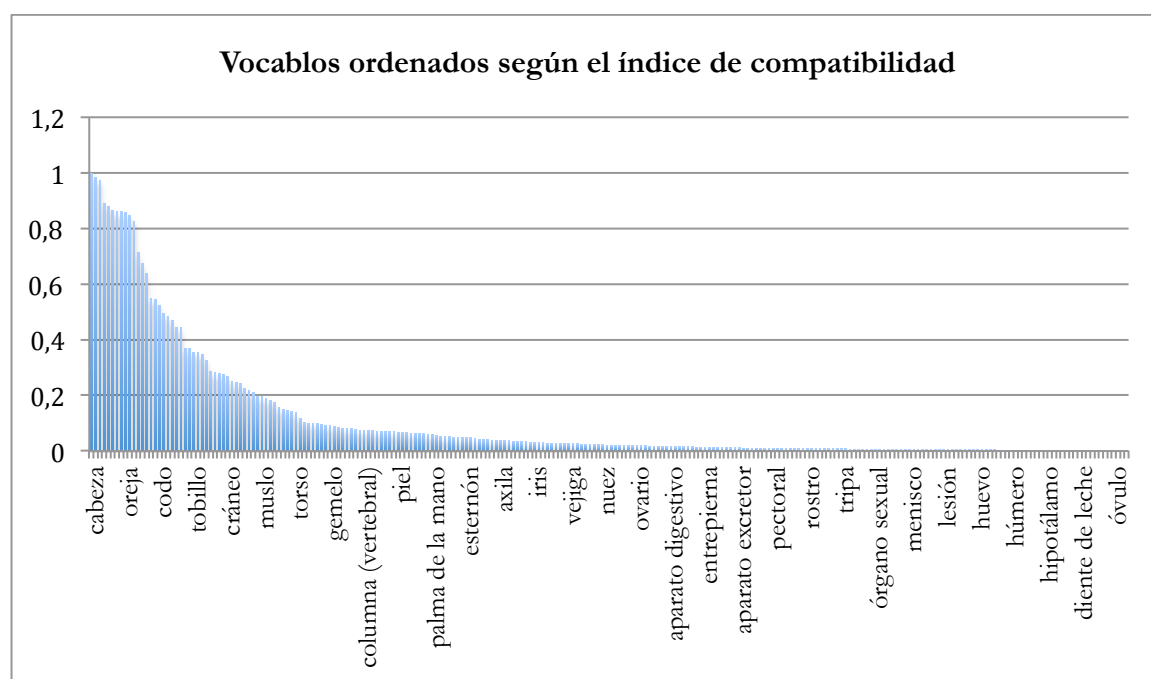


Gráfico 6.1 muestra la gradación de la compatibilidad con respecto al núcleo del

²⁷³ Cabe destacar que si el prototipo estuviera más relacionado con factores sociales o culturales, como pueden ser las costumbres, las respuestas serían diferentes ya que el concepto de prototipo ante esa realidad también lo es. En este sentido, Hernández Muñoz (2006: 145) sugiere que entre los centros de interés tradicionales puede haber alguno que no se corresponda con una categoría prototípica y propone una clasificación de los centros de interés según su estructura interna y niveles de inclusividad (Hernández Muñoz, 2006: 153).

²⁷⁴ Cf. apartado 2.5.2.

centro *Partes del cuerpo humano*. En el caso de tratarse de un gráfico de sedimentación se observaría que llegado a punto no tiene cabida la diferenciación entre un grupo y otro. Por ejemplo, la compatibilidad del vocablo que ocupa el puesto 125, *micHELÍN*, es de 0,0197 y la del último en la clasificación, *ventrículo*, es de 0,00186. Como se muestra en la Tabla 6.1, estos vocablos no forman parte del léxico nuclear o prototípico sino que quedan fuera del nivel de corte establecido a partir del concepto matemático de los conjuntos difusos.

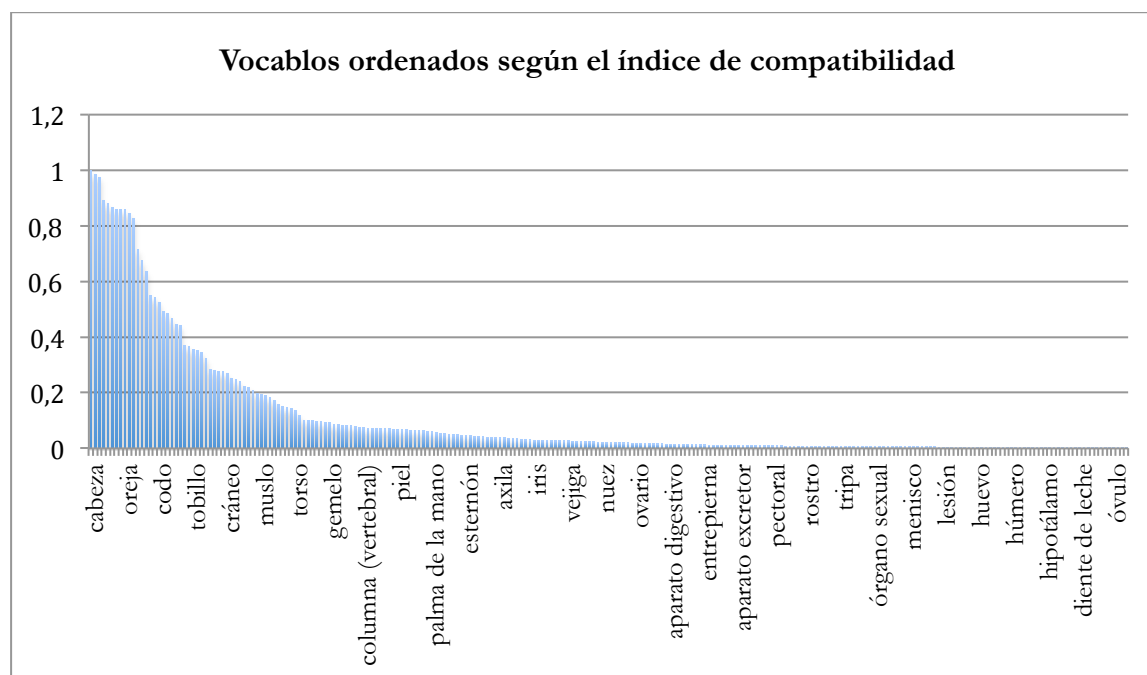


Gráfico 6.1. Vocablos ordenados según el índice de compatibilidad en español. Centro de interés 01. *Partes del cuerpo humano*

Hasta el momento no hay consenso a la hora de seleccionar el vocabulario considerado más disponible. En algunas investigaciones se opta por incluir el índice acumulado de disponibilidad de al menos el 30 % y un índice de disponibilidad léxica mayor de 0,1 (Carcedo González 2001: 62). Otros investigadores, como Bartol Hernández (2001: 227-230) recomiendan usar como criterio de corte únicamente el índice de disponibilidad léxica con un nivel superior o igual al 0,2. En la mayoría de los casos esa selección atiende a fines comparativos de corte psicolingüístico o sociolingüístico y, en raras ocasiones, se ha determinado el léxico relevante.

A partir de los resultados de la prueba de léxico disponible, Ávila Muñoz y Sánchez Sáez (2010, 2011, 2014) consideran que se puede establecer una categorización de las palabras que permita establecer diferentes grados de

accesibilidad. Para ilustrar los preceptos de su teoría, elaboran un cuadro en el que distinguen tres niveles divididos por la línea de seguridad conceptual del centro de interés, la línea de incertidumbre y el límite de categorización conceptual del centro de interés (Ávila Muñoz y Sánchez Sáez, 2010: 59).

La línea de seguridad conceptual delimita aquellos vocablos próximos al núcleo que son compartidos por la mayoría de los informantes y cuentan con un alto índice de disponibilidad léxica. Conforme las palabras se alejan del núcleo se desplazan en la escala de tipicidad hasta llegar a la línea de incertidumbre en la cual la pertenencia de las unidades léxicas al prototipo no es evidente. Una vez rebasadas esas líneas, las palabras son fruto de asociaciones individuales que mantienen en cierto modo una conexión indirecta con el núcleo. No obstante, si la división atiende solo a los criterios de la fórmula de disponibilidad léxica, es decir, número de repeticiones y posición de las palabras, en este último nivel también se registrarían los vocablos técnicos o especializados proporcionados por una minoría de hablantes. En cambio, el índice de compatibilidad tiene en cuenta la relación de la unidad léxica con respecto al núcleo.

Los diferentes niveles expuestos en el Gráfico 6.2 se asemejan a los valores dados a los vocablos para calcular el índice corregido de amplitud léxica (ICoAL). Este índice calcula la amplitud léxica o número total de palabras aportadas por un informante ponderando cada vocablo según su relación con el centro de interés. A modo de recordatorio, las palabras nucleares resultantes de asociaciones directas tienen el valor de 1 (N1), las asociaciones colectivas adquieren un valor de 0,5 (N5), las palabras derivadas 0,25 (N25) y aquellas asociaciones fruto de una asociación individual quedan fuera del prototipo y, por tanto, no adquieren valor alguno (N0). Sin embargo, esta clasificación se basa en criterios puramente semánticos para establecer niveles de tipicidad de las palabras.

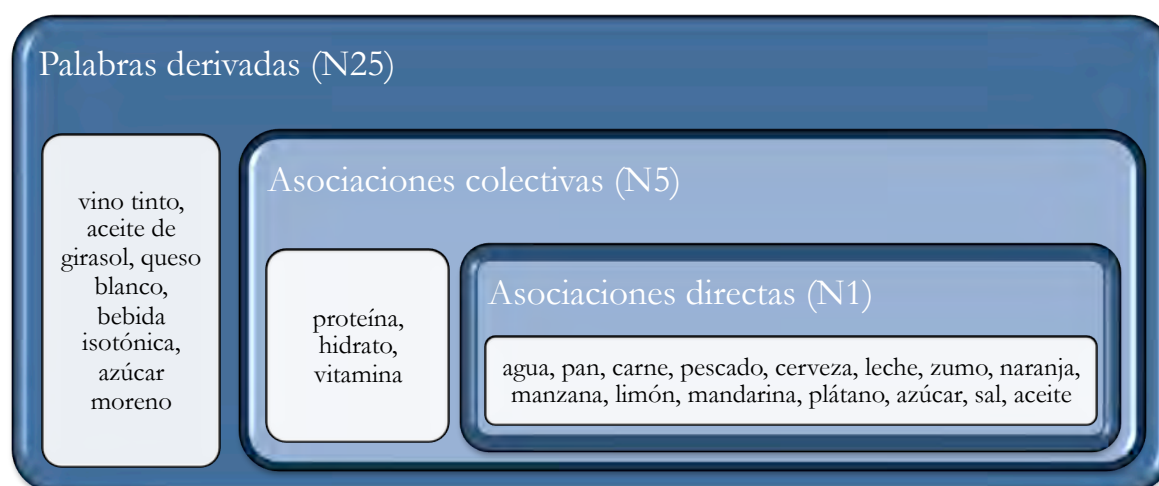


Gráfico 6.2. Niveles de prototipicidad del centro de interés *Comidas y bebidas*²⁷⁵

Según la concepción del modelo anterior, Ávila Muñoz, Sánchez Sáez y Villena Ponsoda aplicaron la teoría de conjuntos difusos, *Fuzzy Expected Value* (FEV), a los estudios de disponibilidad léxica²⁷⁶. A grandes rasgos, su cálculo permite diferenciar en varios niveles la compatibilidad²⁷⁷ del léxico con respecto al centro de interés. Cada nivel recoge un número reducido de vocablos que va aumentando según sea menor el grado de compatibilidad. Además, esos parámetros permiten identificar claramente cuáles son los que cuentan con una alta disponibilidad en los informantes y, por tanto, no deberían faltar en un diccionario.

Este FEV es capaz de proponer un valor de pertenencia que establece un equilibrio entre el número de términos que lo supera y el propio valor. Además, si ponderamos en diferente grado la relevancia del tamaño del conjunto de corte, se pueden establecer grados de restricción en su interior que nos permitirían construir conceptos como «muy representativo» o «poco representativo» (Ávila Muñoz y Sánchez Sáez, 2010: 63).

Los resultados obtenidos permiten diferenciar hasta seis niveles de compatibilidad con el área temática. Cuanto más se acerquen los valores a 1, mayor será su representatividad en el prototipo. El corte en niveles en cada centro varía en función de cada estudio y cada centro y vendrá determinado por el FEV. En lengua española, de los 244 vocablos actualizados en el primer centro de interés, solo 43

²⁷⁵ Este centro de interés recoge más asociaciones derivadas que el resto porque los informantes aportan con frecuencia más datos dentro de una misma categoría. Por ejemplo, se han encontrado las siguientes asociaciones a partir de *queso*: *queso azul*, *queso blanco*, *queso brie*, *queso manchego* y *queso de untar*.

²⁷⁶ Véase Ávila Muñoz y Villena Ponsoda (2010: 55-65), Ávila Muñoz y Sánchez Sáez (2011), Ávila Muñoz y Villena Ponsoda (2012), Ávila Muñoz y Sánchez Sáez (2014: 260-264).

²⁷⁷ En la revisión del modelo matemático del apartado 2.5.2 se determina el índice de compatibilidad.

forman parte de los seis niveles considerados centrales, es decir, un 17,62 %. En ese mismo centro, en inglés han sido seleccionados 26 vocablos de los 115 aportados, correspondiendo a un 22,61 % del total. En francés de las 78 lexías diferentes se han extraído 22, lo que equivale a un 28,21 %.

Estos datos nos ayudarán en la selección del léxico, sobre todo a nivel pedagógico, ya que estas serían las primeras palabras en relación al cuerpo humano que el estudiante debería aprender. A continuación, se presenta a modo de ejemplo los resultados en los tres idiomas para el primer centro de interés. El resto del léxico compatible se encuentra en el anexo VII.

Nivel	Compatibilidad	Proporción	Nº vocablos	Vocablos
1	0,88	1,64 %	4	cabeza, ojo, mano, pie
2	0,82	4,51 %	11	dedo, brazo, nariz, boca, oreja, pierna, pelo
3	0,69	4,92 %	12	uña
4	0,51	6,97 %	17	cara, cuello, ceja, hombro, codo
5	0,33	11,07 %	27	diente, espalda, rodilla, pestaña, pecho, labio, lengua, tobillo, tronco, corazón
6	0,18	17,62 %	43	muñeca, frente, estómago, mejilla, pulmón, cráneo, barriga, cadera, hueso, culo, antebrazo, cerebro, cintura, muslo, riñón, barbilla

Tabla 6.1. Niveles de compatibilidad de los vocablos en español. Centro 01. *Partes del cuerpo humano*

Nivel	Compatibilidad	Proporción	Nº vocablos	Vocablos
1	0,91	4,35 %	5	head, eye, arm, leg, hand
2	0,84	6,095 %	7	nose, mouth
3	0,74	8,7 %	10	finger, ear, hair
4	0,58	11,3 %	13	shoulder, foot, neck
5	0,43	18,26 %	21	face, feet, knee, teeth, nail, toe, heart, elbow
6	0,23	22,61 %	26	lip, tongue, back, eyebrow, chest

Tabla 6.2. Niveles de compatibilidad de los vocablos en inglés. Centro 01. *Human body parts*

Nivel	Compatibilidad	Proporción	Nº vocablos	Vocablos
1	0,89	2,56 %	2	tête, main
2	0,85	6,41 %	5	œil, pied, bras
3	0,76	10,26 %	8	nez, oreille, jambe
4	0,61	14,1 %	11	bouche, cou, yeux
5	0,44	19,23 %	15	doigt, cheveu, cheville, genou
6	0,28	28,21 %	22	ongle, dent, ventre, dos, coude, langue, épaule

Tabla 6.3. Niveles de compatibilidad de los vocablos en francés. Centro 01. *Parties du corps humain*

Del análisis de los vocablos presentes en los tres idiomas se confirma un alto grado de coincidencia. En primer lugar, todas las palabras que forman parte del

vocabulario nuclear en inglés y francés están presentes en la lista en español. Este hecho se explica porque en español existe un número superior de unidades léxicas por tratarse de la lengua materna de los informantes y haber obtenido una media de palabras por informante más elevada que en lengua extranjera. Existen ciertos vocablos presentes solo en una lengua extranjera por dos razones: la traducción requiere una asociación derivada (*toe: dedo del pie*) y el plural es irregular en un solo idioma (*yeux: ojos* y *feet: pies*).

En cuanto a la comparación de los vocablos en lengua extranjera, se observa que en inglés se incorporan cuatro palabras que no aparecen en francés (*chest, eyebrow, face* y *heart*) lo que supone un nivel de coincidencia del 84,61 %. En cuanto al francés, solo incluye dos palabras diferentes al inglés (*cheville* y *ventre*) de forma que presenta una similitud del 90,91 %.

En sintonía con la clasificación realizada por Hernández Muñoz (2006: 153), *Partes del cuerpo humano* es un centro de interés con una estructura interna «bien definida»²⁷⁸ ya que la familiaridad de algunos vocablos hace que existan diferentes grados de tipicidad. De este modo, palabras nucleares como *esófago* no se encuentran dentro de los seis niveles de compatibilidad establecidos porque aunque sean palabras nucleares no son tan familiares o disponibles como *cabeza* o *mano*. Estos resultados son de gran utilidad para establecer criterios de selección del léxico y avanzar en la creación de un diccionario disponible que recoja las palabras más accesibles a los informantes, que será el objetivo último de este capítulo.

6.4. Procesos de reentrada del léxico

En las pruebas de disponibilidad léxica el informante proporciona diferentes vocablos que le vienen a la mente a partir de un estímulo determinado. Hernández Muñoz (2006: 88) considera que la activación del léxico disponible se produce por dos procesos complementarios: la fluencia semántica y, en menor medida, la asociación libre de palabras. En el primer caso, el hablante proporciona palabras

²⁷⁸ Hernández Muñoz (2006) toma como referencia para la clasificación de este centro el estudio realizado por Larrochell, Richard y Soulières (2000) donde se realiza un experimento sobre los índices de medición de las categorías «bien definidas».

relacionadas con un tema o núcleo y, en el segundo caso, el léxico se actualiza en función de las palabras precedentes.

Ávila Muñoz y Sánchez Sáez (2010, 2011 y 2014) parten de la idea que las palabras aportadas en cada centro de interés giran en torno a un núcleo formando una espiral. Según esta hipótesis, las primeras palabras de cada sujeto comienzan en el centro de la espiral y se alejan de él de forma progresiva. No obstante, la forma de esa línea curva puede verse alterada cuando en el camino aparecen palabras que se acercan o alejan de forma considerable al núcleo.

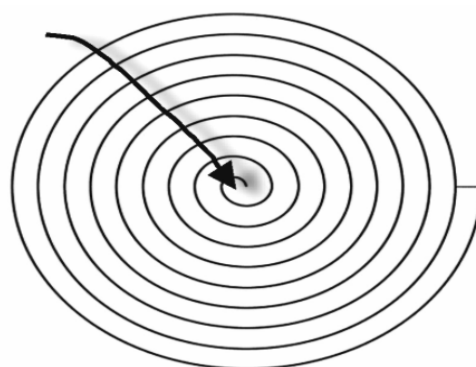


Gráfico 6.3. Puerta de acceso a la red léxica: vocablos muy nucleares (Ávila Muñoz y Sánchez Sáez, 2010: 60)

A partir de los valores calculados mediante el algoritmo de compatibilidad léxica²⁷⁹, Ávila Muñoz y Sánchez Sáez (2010, 2011 y 2014) analizan la variabilidad con respecto al núcleo de las secuencias léxicas aportadas por diferentes informantes. La representación gráfica la realizan incluyendo en el eje de abscisas los vocablos (en el mismo orden aportado por los informantes) y en el eje de ordenadas el índice de compatibilidad de las palabras.

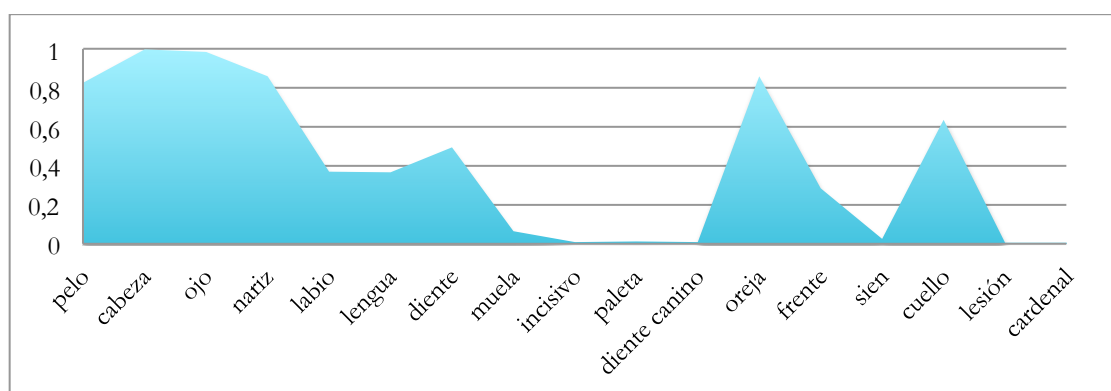


Gráfico 6.4. Muestra de compatibilidad léxica de los 16 primeros vocablos. Informante 109. Centro de interés 01. *Partes del cuerpo humano*

²⁷⁹ Vid. apartado 2.5.2.

Los valores del eje vertical corresponden a la mayor o menor nuclearidad de los vocablos, siendo 1 su máximo valor y 0 el menor. Se observa que los primeros vocablos aportados por los informantes son más compatibles pero, a partir del quinto, con *labio*, se alejan del núcleo considerablemente y la compatibilidad desciende casi a 0. El vocablo *diente* activa en la mente del hablante nombres específicos sobre tipos de dientes: *muela*, *incisivo*, *paleta* y *diente canino*. Este último vocablo activa una palabra más compatible como es el caso de oreja por lo que el vocabulario se acerca de nuevo al núcleo. A continuación, el informante sigue su descripción física ascendiendo en el nivel de compatibilidad y escribe *frente*, pero de nuevo utiliza un vocablo más formal, *sien*, por lo que se aleja del núcleo al cual vuelve a reentrar rápidamente con *cuello*.

La secuencia léxica muestra que el informante comienza aportando palabras que están al mismo nivel con respecto al prototipo (*cabeza*, *ojo* y *nariz*) que además se corresponden con un alto nivel de compatibilidad. En cambio, a partir del vocablo *diente* el sujeto aporta cuatro palabras que guardan una relación de meronimia: *muela*, *incisivo*, *paleta* y *diente canino*. Dentro de esta dimensión vertical de la organización de las palabras, Rosch, Mervis, Gray, Johnson y Boyes-Braem (1976) distinguen tres niveles: nivel supraordinado, nivel de base y nivel subordinado. En este caso el nivel supraordinado viene marcado por el centro de interés *Partes del cuerpo humano* cuya categoría está definida por los vocablos de base (*cabeza*, *ojo*, *nariz*) y los subordinados (*muela*, *incisivo*, *paleta*, *diente canino*).

A los procedimientos de acercamiento y desviación del núcleo prototípico del centro de interés, Ávila Muñoz y Sánchez Sáez los denominan procesos de reentrada. Esos procesos asociativos son analizados de forma individual tomando como referencia las respuestas de un informante en un centro determinado. Tras evaluar el repertorio lingüístico de diferentes hablantes, definen cuatro tipos de patrones de reentrada:

a. Patrón centralizado²⁸⁰. Cuando el informante escribe palabras muy cercanas al núcleo cuentan con un alto índice de disponibilidad (cercano a 1) y dibujan una línea casi horizontal.

²⁸⁰ Todos los ejemplos se han elaborado a partir de los datos del estudio actual.

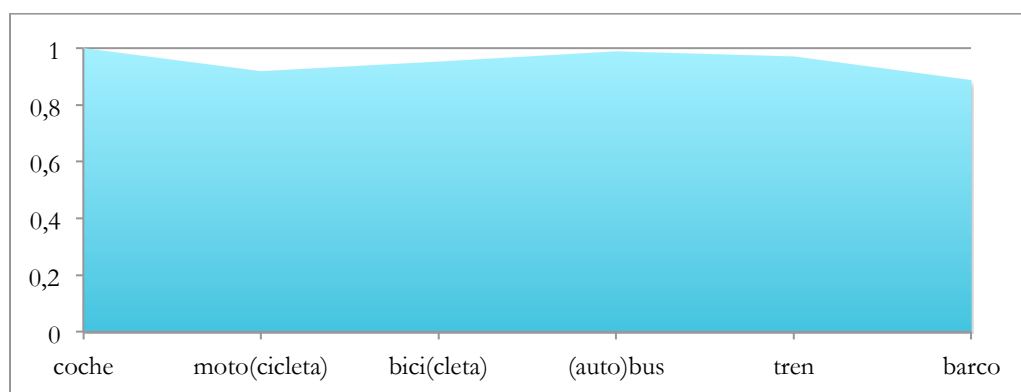


Gráfico 6.5. Muestra de compatibilidad léxica de los 6 primeros vocablos. Patrón centralizado. Informante 157. Centro de interés 06. *Medios de transporte*

b. Patrón de reentrada. El sujeto comienza a escribir palabras centrales pero se desvía hacia lexías menos compatibles. A continuación, vuelve a aportar palabras nucleares y así sucesivamente. Este patrón es el más extendido en todas las respuestas de los informantes.

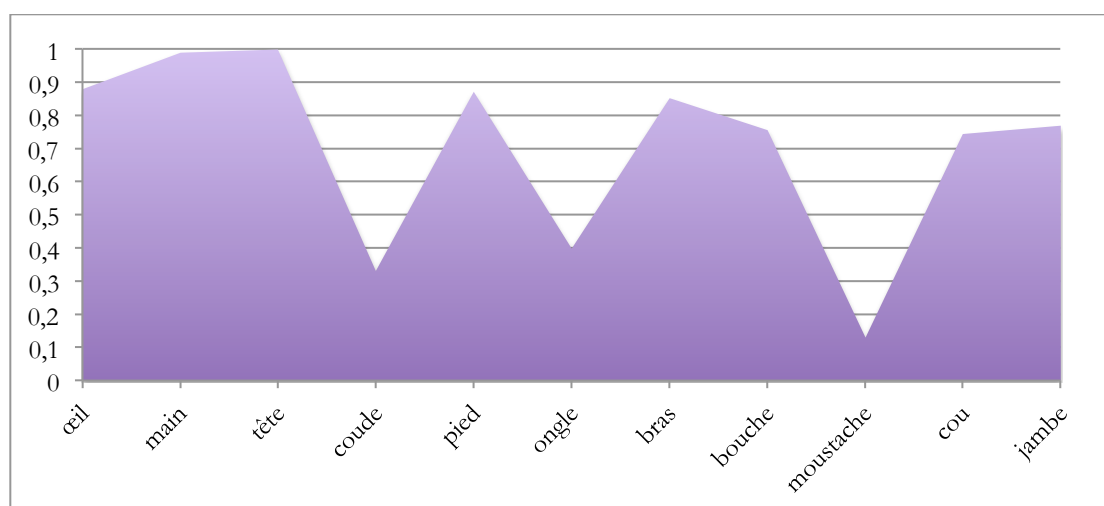


Gráfico 6.6. Muestra de compatibilidad léxica. Patrón de reentrada. Informante 170. Centro de interés 01. *Parties du corps humain*

c. Patrón de desplazamiento y desviación. Se produce cuando el informante comienza a escribir palabras que se desvían del núcleo. Existen diversos factores que influyen en este proceso: asociaciones individuales, repertorio lingüístico específico, derivaciones, etc. Estas palabras suelen quedar fuera de los seis niveles de compatibilidad léxica expuestos anteriormente. No obstante, dentro de ese desplazamiento también se observan palabras que vuelven a acercar el vocabulario al núcleo.

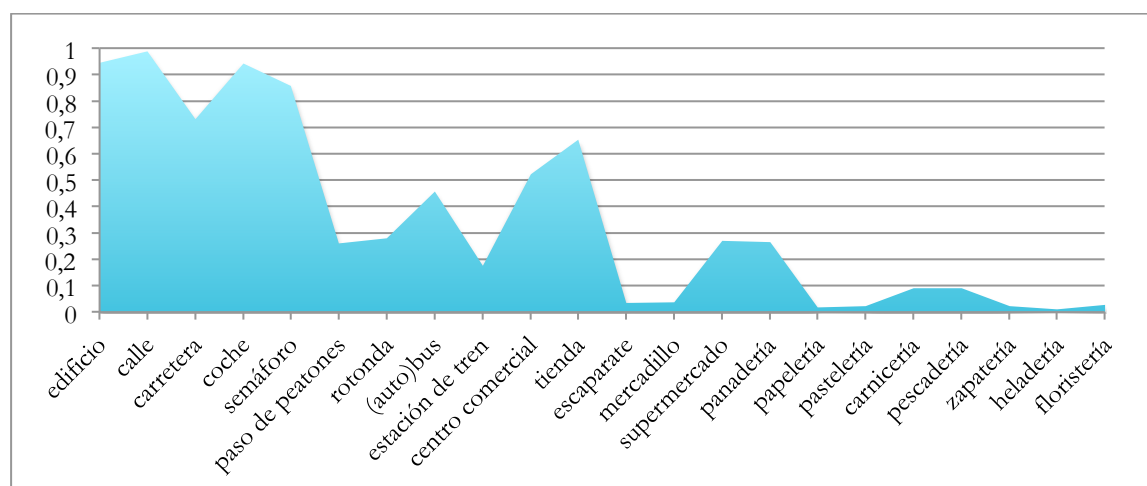


Gráfico 6.7. Muestra de compatibilidad léxica. Patrón de desplazamiento y desviación. Informante 022. Centro de interés 04. *La escuela: muebles y materiales*

d. Patrón de desviación sistemática. Se produce cuando el sujeto escribe palabras menos compatibles de forma progresiva. En este modelo no se producen procesos de reentrada al núcleo por lo que es muy poco frecuente. En el ejemplo se han incluido solo cuatro vocablos porque casi ningún informante presenta este patrón. Tras escribir la palabra *net* el hablante aporta otra palabra que lo lleva a volver al núcleo confirmando que lo más frecuente son los procesos de reentrada.

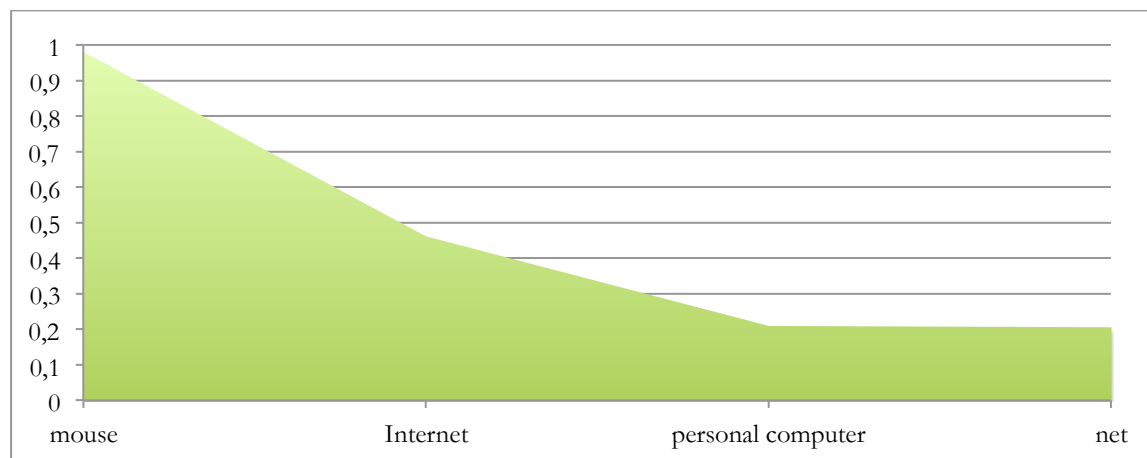


Gráfico 6.8. Muestra de compatibilidad léxica con los vocablos en posición 2, 3, 4 y 5. Patrón de desviación sistemática. Informante 119. Centro de interés 09. *Computer and Internet*

En el presente estudio, esos patrones dominantes en sentido estricto han sido difíciles de identificar si se toma como referencia el conjunto de palabras de un informante en un centro concreto. Por ello, en el caso del patrón centralizado y el patrón de desviación sistemática se han seleccionado solo algunos vocablos que ilustran ese comportamiento. Esas muestras sirven para explicar comportamientos concretos de los informantes en secuencias de palabras dadas. Por ejemplo, en el

Gráfico 6.9, se podría considerar que desde la primera palabra, *oreja* hasta la octava *espalda* se observa un patrón centralizado compuesto por palabras con un alto índice de compatibilidad. A continuación, se produce un patrón de desviación sistemática que va desde el vocablo *boca* hasta *pecho*.

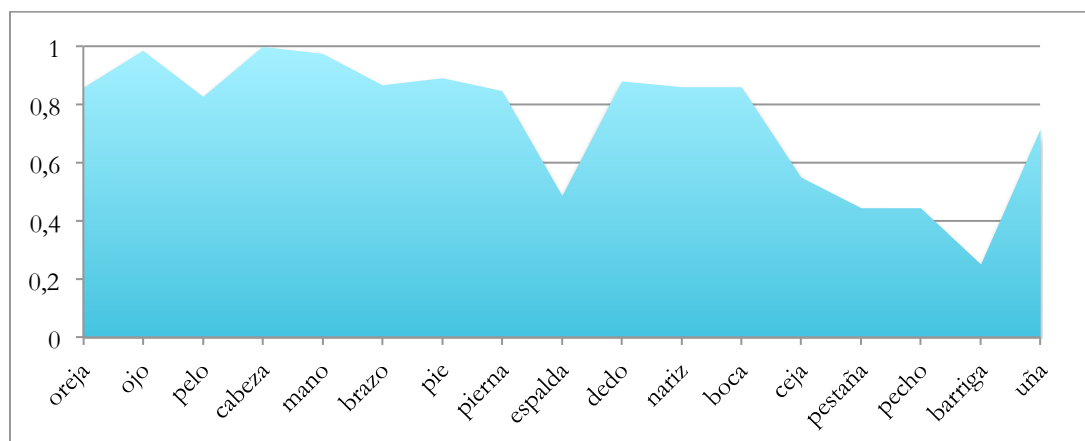


Gráfico 6.9. Muestra de compatibilidad léxica. Informante 062. Centro de interés 01. *Partes del cuerpo humano*

6.4.1. Modelos de reentrada según la ontogenia de los centros de interés

De la cuantificación de la disponibilidad léxica en los centros de interés se desprende una diferente estructura según su productividad, heterogeneidad y densidad. Es frecuente encontrar en los estudios sobre léxico disponible un apartado de análisis comparativos de los vocablos y palabras aportados en los centros de interés seleccionados (Ávila Muñoz, 2010; Sánchez-Saus Laserna, 2011; Bellón Fernández, 2011; entre otros). Asimismo, tomando como referencia esos dos datos se ha tratado de analizar la densidad de los centros a través del índice de cohesión. Se podría decir que con cada uno de los resultados (número de palabras, número de vocablos e índice de cohesión) el investigador sitúa en una escala los centros mostrándolos en una clasificación por rangos.

Gougenheim *et al.* (1964: 189-190) clasifican los centros por orden de dispersión creciente, es decir, según el número de vocablos aportados en cada centro de interés en los diferentes departamentos. Los investigadores franceses distinguen entre dos tipos de centros: aquellos que no llegan a los 400 vocablos y los que superan este umbral. Asimismo, asocian el grado de disponibilidad (basado en la

frecuencia) de la palabra más «disponible»²⁸¹ con la dispersión o número de vocablos de forma que en los centros con más de 400 palabras diferentes el grado de disponibilidad es menor debido a la falta de homogeneidad. Concluyen otorgando mayor relevancia para el estudio a los centros que pertenecen a la primera categoría, razón por la cual analizan más en profundidad *Les vêtements* y *Les parties du corps*²⁸². Los resultados de la prueba de Marne dan lugar a la siguiente clasificación:

Centres d'intérêt	Dispersion	Degré de disponibilité du premier mot
1. Les vêtements	156	93
2. Les parties du corps	177	86
3. Les animaux	220	89
4. Les meubles	244	98
5. Les objets placés sur la table	255	100
6. Aliments et boissons	260	89
7. L'école	280	87
8. L'école	296	90
9. La maison	333	91
10. La cuisine	341	94
11. Chauffage et éclairage	369	82
12. Jeux et distractions	414	71
13. Le village	429	75
14. Les métiers	434	60
15. Les travaux des champs et du jardinage	449	64
16. La ville	473	72

Tabla 6.4. Clasificación de los centros según el número de vocablos (Gougenheim *et al.*, 1964: 189-190)

Desde que Gougenheim *et al.* (1964) destacaran la similitud entre el comportamiento de los centros de interés de las cuatro encuestas llevadas a cabo y la clasificaran según la homogeneidad, se han propuesto diferentes tipos de clasificación (Hernández Muñoz, 2006; Sánchez-Saus Laserna, 2011; Tomé Cornejo, 2011). Hernández Muñoz (2006) dedica un apartado de su tesis doctoral a la categorización de los centros de interés que concluye con un bosquejo de la taxonomía de los centros.

En primer lugar, la autora distingue dos tipos de centros según los niveles de inclusividad: centros inclusivos y relacionales. Los primeros están formados por elementos que mantienen una asociación directa con la categoría y donde aplica la ecuación «x es Y». Por ejemplo, «la mano es una parte del cuerpo». En cambio, en los centros relacionales los elementos no se relacionan de forma tan evidente con la

²⁸¹ Aquella que se encuentra con más frecuencia en los listados.

²⁸² « Seuls donc les résultats de la première catégorie (centres d'intérêt homogènes) sont à retenir pour l'étude du rapport envisagé » (Gougenheim *et al.*, 1964: 190).

categoría. En este caso, enuncia la proposición «x es parte de Y» y, más precisamente, «x está relacionado con Y» o «x tiene algún tipo de relación con Y». Por ejemplo, «las Matemáticas están relacionadas con la escuela» (Hernández Muñoz, 2006: 151-152).

A continuación, Hernández Muñoz (2006) analiza en detalle la estructura interna de los centros de interés en su dimensión horizontal y distingue cuatro tipos: natural, bien definida, *ad hoc* y radial. Por categoría natural entiende aquellas que presentan las características clásicas de una categoría semántica justificada deontológicamente por la teoría de los prototipos, la teoría de los ejemplares o la teoría del conocimiento (Hernández Muñoz, 2006: 135). La categoría bien definida se corresponde con los centros que tienen una estructura clara y cerrada. La categoría *ad hoc* está compuesta por elementos que se organizan en torno a una finalidad o propósito, como por ejemplo: enumerar los ingredientes para hacer una paella. Por último, la categoría radial se configura en forma de red en la que existen subcategorías compuesta por diferentes elementos periféricos.

La clasificación de Hernández Muñoz organizaría de la siguiente forma los centros de interés del presente estudio:

CI	Estructura interna				Niveles de inclusividad	
	natural	bien definida	ad hoc	radial	inclusiva	relacional
1. Partes del cuerpo		X			X	
2. La ropa	X				X	
3. Comidas y bebidas	X				X	
4. La escuela: muebles y materiales ²⁸³				X	X	
5. La ciudad				X		X
6. Medios de transporte	X				X	
7. Juegos y distracciones				X	X	
8. Profesiones y oficios	X				X	
9. Ordenadores e internet				X		X

Tabla 6.5. Clasificación de los centros de interés del presente estudio según la estructura de organización interna y sus niveles de inclusividad²⁸⁴

Esta clasificación incluye ocho centros de los dieciséis tradicionales y uno, *Ordenadores e internet*, cuyas características no han sido descritas previamente. Los análisis llevados a cabo en el plano cualitativo y cuantitativo indican que este centro posee una estructura interna radial que podría ser considerada más específicamente

²⁸³ A pesar de que la autora incluye *La escuela* en la categoría relacional, aclara en una nota a pie de página que se debe a que en la prueba realizada en Castilla La Mancha no limitaron el ámbito léxico a *muebles y materiales* (Hernández Muñoz, 2006: 151). En tal caso, se consideraría una categoría inclusiva, tal y como ocurre en este estudio.

²⁸⁴ Clasificación realizada en base a la propuesta por Hernández Muñoz (2006: 153).

dentro de la categoría de esquema porque en ocasiones el informante describe espacialmente los elementos que componen ese centro de interés. En cuanto a los niveles de inclusividad, se constata que es un centro heterogéneo que presenta una categoría relacional cuyas palabras están relacionadas con el núcleo sin necesidad de que pertenezcan estrictamente a un prototipo. Para justificar esa categorización, en el apartado 6.4.5 se analizan los procesos de reentrada y en el apartado 6.5.2.3 las relaciones entre las palabras de ese centro.

Tomé Cornejo (2011) introduce una nueva categoría denominada *esquema* basada en la distinción de Richards (1969). Los elementos están relacionados con los espacios del mundo real que el informante puede describir desde diferentes ópticas: objetos, sentimientos, personas, etc. Los centros organizados según esta categoría serían: *La cocina y sus utensilios*, *La escuela*, *La ciudad*, *El campo* y *Juegos y distracciones*.

Tomando como referencia la taxonomía de Hernández Muñoz (2006) y Tomé Cornejo (2011), Sánchez-Saus Laserna (2011) realiza una clasificación de los centros de interés en la que distingue entre categorías naturales, categorías radiales y esquema. Dentro de esta última categoría identifica una subcategoría, el guión, y lo atribuye al centro denominado *Acciones y actividades habituales* por estar más relacionado a lo eventual y a la secuenciación de acciones (Sánchez-Saus Laserna, 2011: 207).

López González (2014: 49) muestra una clasificación de los centros de interés que recoge las aportaciones de las tres investigadoras anteriores. No obstante, la clasificación de la categoría radial y el esquema coinciden en todos los centros de interés por lo que no tendría sentido tal distinción en los centros de interés tradicionales. Asimismo, asocia el centro de interés *El cuerpo humano* tanto a la categoría natural como a la bien definida.

Al igual que ocurre en la investigación llevada a cabo por Sánchez-Saus Laserna (2011), los criterios de edición han permitido la inclusión de todo tipo de palabras que el informante ha relacionado con el centro de interés difuminando los límites de las categorías naturales y asemejándolas a las categorías radiales. Por esta razón, aunque *Partes del cuerpo humano* sea una categoría bien definida, esa propiedad no se refleja en los resultados ya que se han incluido vocablos en ese centro de interés como *hidrato* o *dieta* que no son en sí mismos partes del cuerpo sino que son

fruto de asociaciones indirectas. A pesar de que Sánchez-Saus Laserna (2006: 206) considera que los criterios de edición abiertos eliminan en parte las características de las categorías naturales de los centros de interés, ve necesario hacer la distinción entre los centros radiales y los centros naturales, además de incorporar un tercer tipo denominado esquema.

La propuesta de clasificación de los centros que se propone a continuación se basa en la organización de las palabras en la mente distinguiendo entre centros prototípicos y centros relacionales. Los primeros hacen referencia a centros homogéneos cuyo léxico nuclear se encuentra mejor definido. En cambio, los segundos son menos compactos y las relaciones entre las palabras están menos organizadas. No se trata de una dicotomía entre centros prototípicos y centros relacionales sino más bien en localizar los dos extremos de un continuo en el que se sitúan los centros de interés.

A grandes rasgos, los centros prototípicos suelen coincidir con los centros naturales y los centros relacionales con los centros radiales pero la distinción se ha realizado adecuando la teoría a los resultados obtenidos en este estudio. No obstante, existen ciertas divergencias ya que se ha pretendido el esbozo de una taxonomía que refleje la estructura interna de los centros de interés atendiendo a tres criterios: la centralidad, la cohesión y la conectividad de los vocablos. Por tanto, la clasificación está influida por el número de informantes y sus características y, por supuesto, por el tipo de estímulo que provoca una estructura u otra en cada centro de interés.

Para la clasificación según rangos de los centros en función de su estructura interna se ha utilizado el índice estructural definido como la suma del índice de cohesión, el índice de centralidad y el índice de conectividad. El índice de cohesión muestra la relación entre la media de palabras por informantes y el número total de vocablos. Este índice oscila de 0 a 1 de forma que cuanto más se acerque a 1 más homogéneo será el vocabulario y viceversa, cuanto más se acerquen a 0 más heterogéneas serán las respuestas de los informantes.

Tomando como referencia el número de vocablos nucleares se ha generado el índice de centralidad que muestra la cantidad de vocablos nucleares en cada centro con respecto al total de vocablos. Este indicador se obtiene al dividir el número total

de vocablos por el número de vocablos considerados centrales o nucleares. El índice oscila entre un máximo de 1, correspondiente a un número total de palabras y vocablos nucleares idéntico, y un mínimo de 0, en el supuesto que no existiese ningún vocablo nuclear.

$$\text{Índice de centralidad}^{285} = \frac{\sum V}{\sum VN}$$

En el índice de conectividad se han tenido en cuenta el número de conexiones con independencia de su peso. Obviamente, si se consideran todas las conexiones y se suman los pesos, el resultado expresará que el total de palabras está relacionado, ya que se considera la relación unidireccional de una sola palabra por vocablo y eso siempre se produce. Por ello, se ha considerado el número de conexiones entendiendo que cuanto menor sea el número de conexiones con respecto al número de palabras aportadas, mayor será la fuerza de la conexión entre esas palabras. El índice de conexión se obtiene al dividir el número total de conexiones por el número de palabras y restarle 1. En la primera operación se obtiene el número de relaciones que se han producido por cada palabra que el informante ha aportado. Dentro de las relaciones entre palabras se han sumado todas aquellas diferentes, por lo que cuanto menor sea el número de relaciones mayor será su peso o fuerza. El índice oscila entre 0 y 1, donde 0 equivaldría a ninguna relación por palabra (lo cual es imposible) y 1 a que todas las palabras mantienen relaciones diferentes. Para poder realizar una comparativa con el índice de centralidad y de cohesión en los que se muestra la homogeneidad del núcleo, se ha restado 1 al resultado de forma que los datos que se acercan a 1 muestran conexiones con mayor peso en las palabras y los resultados que se acercan a 0 relaciones más difusas.

$$\text{Índice de conexión}^{286} = 1 - \left(\frac{\sum C}{\sum P} \right)$$

La Tabla 6.6 muestra los resultados ordenados según el rango que ocupan los centros de interés en cada índice. Se han destacado en sombreado los tres primeros centros y los tres últimos de forma que se establecen tres intervalos según su homogeneidad. Se constata la similitud entre las respuestas de un mismo centro en

²⁸⁵ Donde $\sum V$ hace referencia a la suma de vocablos y $\sum VN$ a la suma de vocablos nucleares.

²⁸⁶ Donde $\sum C$ se refiere a la suma de conexiones diferentes y P al número de palabras.

los tres índices. Así, *Partes del cuerpo humano* ocupa el primer lugar en cuanto al índice de cohesión, índice de centralidad y en el índice estructural pero ocupa la segunda posición en el índice de conectividad (con un índice de 0,548 muy similar al centro con mayor conectividad, *Medios de transporte*, que cuenta con un índice de 0,554).

CI	Índice de cohesión	CI	Índice de centralidad	CI	Índice de conectividad	CI	Índice estructural
1	0,108	1	0,176	6	0,554	1	0,832
2	0,092	2	0,172	1	0,548	2	0,791
6	0,081	6	0,147	2	0,526	6	0,782
4	0,055	8	0,146	4	0,471	4	0,655
3	0,05	5	0,139	3	0,330	3	0,518
5	0,045	3	0,138	5	0,319 ²⁸⁷	5	0,503
8	0,044	7	0,136	7	0,319	7	0,493
7	0,038	4	0,128	9	0,314	8	0,475
9	0,035	9	0,124	8	0,284	9	0,473

Tabla 6.6. Indicadores de la estructura de los centros de interés en español

Los centros *Partes del cuerpo humano* (1), *La ropa* (2) y *Medios de transporte* (6) se sitúan en las tres primeras posiciones del rango. Según la clasificación de Hernández Muñoz (2006), el primer centro es considerado una categoría bien definida y los otros dos categorías naturales según su estructura interna pero los tres son inclusivos según su nivel de inclusividad. *Medios de transporte* ocupa el tercer puesto en cuanto a índice de cohesión e índice de conectividad. Se trata del centro con el menor número de vocablos, lo que indica que el núcleo está mejor definido. En cambio, las relaciones entre las palabras son más diversas que en los dos centros primeros por lo que ocupa el quinto lugar.

En el nivel intermedio se sitúan los centros *La escuela: muebles y materiales*, *Comidas y bebidas* y *La ciudad*. La escuela es un centro clasificado en Hernández Muñoz (2006) con una categoría radial. De hecho en el índice de nuclearidad ocupa el penúltimo lugar. En cambio, en el índice de cohesión y el índice de conectividad ocupa el cuarto lugar. Ello indica que aunque el núcleo sea algo más amplio las palabras están cohesionadas y conectadas.

Comidas y bebidas presenta un núcleo con una proporción de palabras superior por lo que ocupa el sexto lugar. No obstante, en el índice de cohesión se sitúa en el quinto lugar y en el índice de conectividad en el tercero. Se trata de un centro de

²⁸⁷ Para ordenar el centro 5 y 7 se han tenido en cuenta las siguientes cifras decimales. En el caso del centro 5 el resultado es 0,319051 y en el del centro 7 es 0,318750.

interés con una estructura bien definida y con unas relaciones claras aunque también sea un centro donde, en ocasiones, los informantes aportan nombres específicos de comidas o realizan asociaciones individuales que llevan a desviarse del núcleo.

La ciudad es un centro clasificado como radial y de categoría relacional por Hernández Muñoz (2006). Ocupa el quinto lugar en el índice de nuclearidad pero se sitúa en el sexto en cuanto al índice de cohesión y el séptimo en el índice de conectividad. Se trata de un centro muy productivo y con una gran riqueza léxica ya que es un campo abierto en el que los límites del prototipo son bastante difusos.

En el extremo inferior se sitúan los centros, *Juegos y distracciones*, *Profesiones y Oficios* y *Ordenadores e internet*. El primero es un centro claramente relacional en el que los límites no están tan bien definidos debido a que existen muchas formas de pasar el tiempo libre y distraerse que dependen directamente de los informantes y de sus características (por ejemplo, el golf aparecerá con más frecuencia en los informantes con un poder adquisitivo mayor). En el índice de centralidad ocupa el séptimo lugar, en el índice de cohesión el octavo y en el de conectividad de los vocablos el sexto lugar. La variedad de respuestas en este centro hace que las relaciones entre las palabras sean más diversas y el prototipo esté más disperso.

Profesiones y oficios es un centro en el que se han producido una serie de asociaciones indirectas que llevan a aportar más vocablos diferentes, lo que sitúa en el séptimo puesto del rango según el índice de cohesión. En cambio, la proporción de vocablos nucleares con respecto al total de vocablos es mayor y se sitúa en el cuarto puesto, lo que indica que es un centro cuyo núcleo está bien definido. El hecho de que los estudiantes sean de postgrado y se hayan incorporado o lo vayan a hacer en breve, quizás ha llevado a un mayor conocimiento de distintas profesiones y a asociaciones indirectas e incluso individuales.

Ordenadores e internet se sitúa en el índice de cohesión, de centralidad y estructural en el último rango. Tan solo sube un puesto en el índice de conectividad gracias a relaciones entre palabras muy frecuentes (por ejemplo *teclado* y *ratón* han sido aportadas en uno u otro orden por 75 informantes). El ordenador podría considerarse un centro inclusivo con una estructura natural si el informante se limitara a enumerar sus partes, ya que si se abre el campo a nombres de programas

informáticos o software la lista sería encomiable. Asimismo, los límites de internet son muy amplios lo que lleva a que se trate de un centro relacional donde se incluyan relaciones de todo tipo (nombres propios de páginas webs, nombres de programas, etc.).

En inglés, la distribución de los centros según rango es semejante pero está mejor definida. Se han marcado los centros que coinciden en los cuatro índices en negrita y se observa que repiten posición cinco centros en todos sus índices. En los tres primeros lugares se sitúan los mismos centros que en español (*Human body parts*, *Clothes*, *Means of transport*). A continuación, se incluyen los centros intermedios en el que se cambiado el centro número cinco, *La ciudad*, por el centro número ocho, *Profesiones y Oficios*. Esto puede deberse a que en el aprendizaje de una lengua extranjera existe más vocabulario relacionado con La ciudad que con las Profesiones y Oficios, razón por la cual los informantes han aportado más variedad de unidades léxicas.

CI	Índice de cohesión	CI	Índice de centralidad	CI	índice de conectividad	CI	Índice estructural
1	0,149	1	0,226	1	0,586	1	0,961
2	0,092	2	0,185	2	0,508	2	0,785
6	0,068	6	0,174	6	0,372	6	0,761
3	0,067	4	0,131	4	0,469	4	0,658
4	0,058	3	0,157	3	0,303	3	0,614
8	0,046	8	0,158	8	0,536	8	0,518
5	0,046	7	0,145	7	0,322	7	0,508
7	0,041	5	0,157	5	0,315	5	0,506
9	0,033	9	0,117	9	0,280	9	0,430

Tabla 6.7. Indicadores de la estructura de los centros de interés en inglés

En francés, los tres primeros centros coinciden con la clasificación en español y en inglés (*Parties du corps humain*, *Moyens de transport* y *Les vêtements*). Los tres que se sitúan en la parte central de la escala son los mismos que en español. No obstante, se han marcado en violeta los centros que se sitúan en su gran parte en el subgrupo formado por los centros con índices menores. Cabe destacar que el centro nueve, *Ordinateurs et Internet*, al igual que en español, no se sitúa en el último rango según conectividad quizás debido a la existencia de conexiones muy frecuentes y que este centro no cuenta con la misma riqueza léxica que en inglés.

En cambio, el quinto centro presenta una estructura más homogénea en francés que en inglés aunque ocupa el último rango según conectividad. Como se comentó anteriormente, los índices están influenciados por el número de informantes y en este caso el número de informantes es considerablemente menor por lo que quizás los datos no se encuentren tan bien definidos como en español o inglés. Por ello, la comparación se hace en función de los rangos que ocupan los centros en cada índice con independencia del número exacto en cada caso.

CI	índice cohesión	CI	Índice de centralidad	CI	índice de conectividad	CI	Índice estructural
1	0,261	6	0,302	6	0,353	6	0,879
6	0,224	1	0,282	1	0,265	1	0,807
2	0,217	2	0,275	2	0,255	2	0,747
5	0,137	5	0,237	4	0,201	4	0,559
4	0,136	8	0,235	7	0,152	5	0,495
3	0,130	3	0,223	9	0,152	3	0,494
8	0,120	4	0,221	3	0,142	9	0,491
9	0,118	9	0,220	8	0,127	8	0,482
7	0,107	7	0,2	5	0,121	7	0,460

Tabla 6.8. Indicadores de la estructura de los centros de interés en francés

El objetivo de este apartado ha sido demostrar el diferente comportamiento de los centros de interés y establecer una medida de valoración de sus características. De la comparación de las similitudes y divergencias en la clasificación en cada idioma, se pueden extrapolar algunos resultados generales compartidos en gran medida en los tres estudios. A modo de ejemplo, se presenta el continuo que formarían los centros ordenados según el rango español. En el extremo izquierdo se sitúa el centro más prototípico, *Partes del cuerpo humano*, y en el derecho el centro menos prototípico o relacional, es decir, con una estructura más abierta a relaciones y asociaciones diferentes, *Ordenadores e internet*.

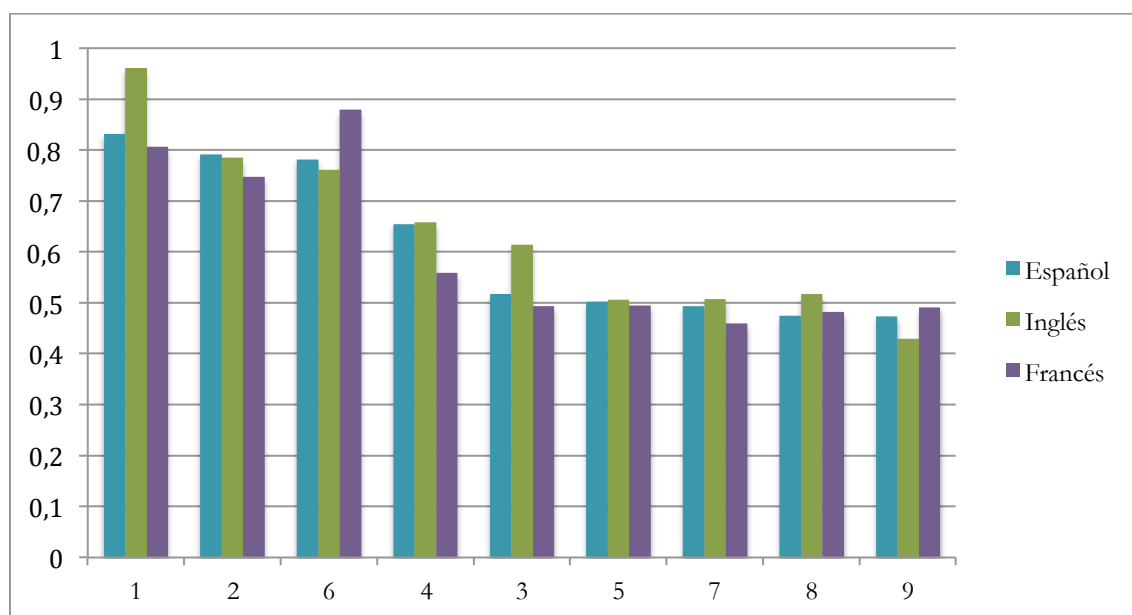


Gráfico 6.10. Escala de prototipicidad de los centros de interés

En el Gráfico 6.10 se destacan los tres primeros centros por su prototipicidad referidos a las *Partes del cuerpo*, *La ropa* y *Los medios de transporte*. A continuación, le sigue un centro que podría considerarse, *La escuela: muebles y materiales*. Los cinco últimos centros de interés presentan resultados similares en los tres idiomas a excepción del centro de *Comidas y bebidas* que tiene en inglés un índice ligeramente superior a los otros dos idiomas pero por debajo del centro que le precede en la clasificación.

A la vista de los resultados anteriores se propone la siguiente clasificación:

Número	Centro	Centros prototípicos	Centros relacionales
1	Partes del cuerpo	X	
2	La ropa	X	
3	Comidas y bebidas		X
4	La escuela: muebles y materiales		X
5	La ciudad		X
6	Medios de transporte	X	
7	Juegos y distracciones		X
8	Profesiones y oficios		X
9	Ordenadores e internet		X

Tabla 6.9. Propuesta de clasificación de los centros de interés

6.4.2. Modelo de reentrada de los centros de interés prototípicos

Los centros de interés considerados prototípicos presentan tres características comunes: un vocabulario homogéneo, un léxico nuclear concreto y unas relaciones entre las palabras bien definidas. Quizás la primera característica sea la más

importante de todas ya que condiciona a las otras dos: si el vocabulario es más homogéneo el léxico nuclear está más concentrado y las relaciones entre unidades léxicas son más claras ya que no dan lugar a asociaciones individuales. Los tres centros clasificados como prototípicos son, en los tres idiomas, *Partes del cuerpo humano*, *La ropa* y *Medios de transporte*.

En este apartado se presentan los procesos de reentrada de los centros de interés prototípicos tomando como referencia el centro denominado *Partes del cuerpo humano* por tratarse de un centro cuyos límites categoriales se encuentran mejor definidos y delimitados. En primer lugar, se presentan los análisis de los informantes que han aportado pocas palabras en los tres idiomas para mostrar el modelo común que presenta ese tipo de secuencias léxicas: el modelo de reentrada leve. En segundo lugar, se muestran las respuestas de hablantes que se sitúan en torno a la media y muestran un modelo de reentrada moderado. En tercer lugar, se incluyen los modelos de los sujetos con mayor capacidad léxica que se situarían en el extremo de este continuo produciendo un modelo de reentrada pronunciado.

6.4.2.1. Modelo de reentrada leve

Se produce cuando los procesos de reentrada son escasos e implican vocablos con diferencias poco significativas en el índice de compatibilidad. En general, todas las palabras aportadas por el informante suelen situarse por encima de la media del índice de compatibilidad de ese centro de interés. En estos casos, el léxico suele ser reducido, lo que implica que la mayoría de las palabras de un sujeto tengan un alto índice de compatibilidad. Por tanto, este modelo podría atribuirse a los hablantes cuyas palabras se acercan más al núcleo, pero con una productividad y originalidad limitada.

Para ilustrar el modelo, se ha seleccionado al informante número 60 por ser el que se sitúa en último lugar en cuanto al número de palabras en el centro número 1. Teniendo en cuenta que la media de compatibilidad de los vocablos para el centro 1 es de 0,112 se constata que todos los vocablos superan ampliamente ese umbral. El proceso de reentrada se observa entre los vocablos en las posiciones 7, 8 y 9. El vocablo *uña* (posición 7) con un índice de compatibilidad de 0,715 lleva a un

descenso al escribir *barriga* (posición 8) con un índice de compatibilidad de 0,251. A continuación, el informante escribe en la posición la palabra *mano*. El vocablo más descentralizado en esta lista ocupa el puesto número 34, según el orden de palabras marcado por el índice de compatibilidad, y forma parte del sexto nivel de distinción según la nuclearidad de las palabras²⁸⁸.

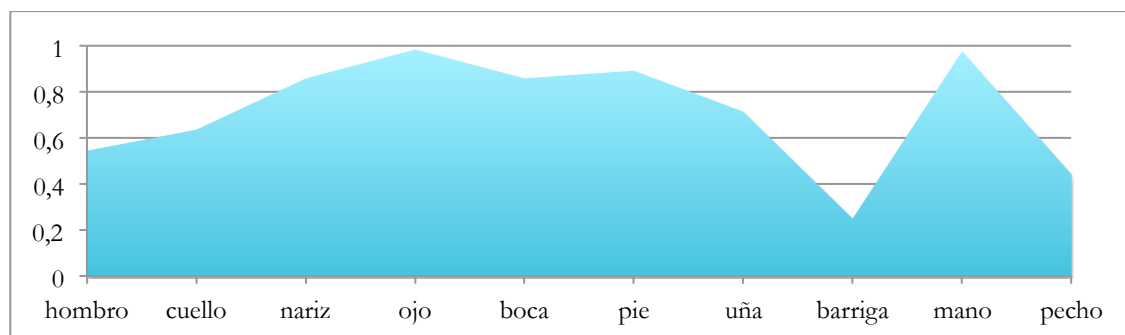


Gráfico 6.11. Muestra de compatibilidad léxica en español. Modelo de reentrada leve. Informante 060. Centro de interés 01. *Partes del cuerpo humano*

En inglés, se ha seleccionado también a uno de los informantes que menos palabras ha aportado, concretamente el número 145. En este centro de interés la media de compatibilidad es de 0,176 por lo que se observa que todos los vocablos, a excepción del segundo superan ese umbral. Esa palabra, *heart*, se sitúa en torno a la media con un índice de compatibilidad de 0,1724. Al igual que en español, todas las unidades léxicas están indexadas como nucleares, incluida *heart*, que forma parte del quinto nivel. Se observa una relación de meronimia entre *hand* y *finger* que en el caso del español se producía entre *pie* y *uña*.

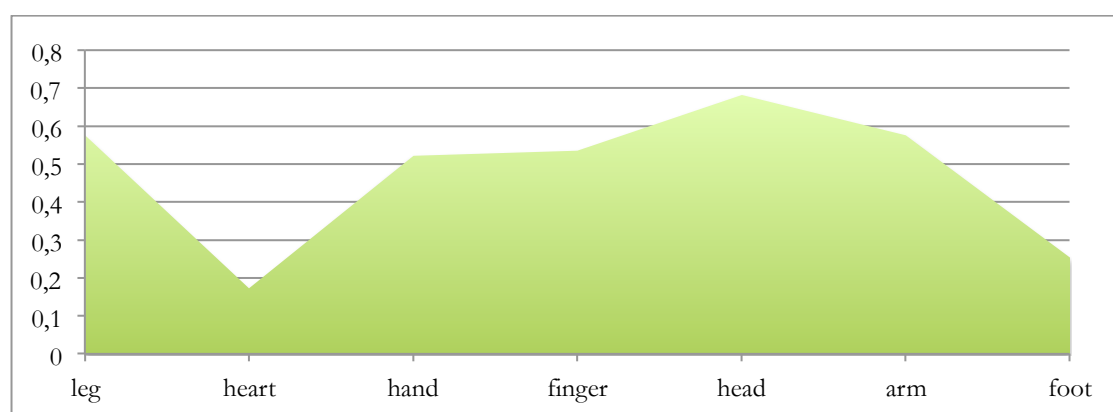


Gráfico 6.12. Muestra de compatibilidad léxica en inglés. Modelo de reentrada leve. Informante 145. Centro de interés 01. *Human body parts*

²⁸⁸ Vid. Anexo VIII para consulta el listado de palabras nucleares.

En francés, se han representado las respuestas del informante 115 que ha aportado 12 palabras. El vocablo que menos índice presenta es *langue* con 0,296 pero se sitúa por encima de la media de compatibilidad con 0,2249. Todas las palabras forman parte del vocabulario nuclear, incluido el vocablo *langue* que pertenece al nivel 6 de compatibilidad. Al igual que en español, donde es muy frecuente la relación *mano* y *pie*, los dos primeros vocablos del informante han sido *main* y *pied*. A continuación, le sigue su holónimo *jambe*.

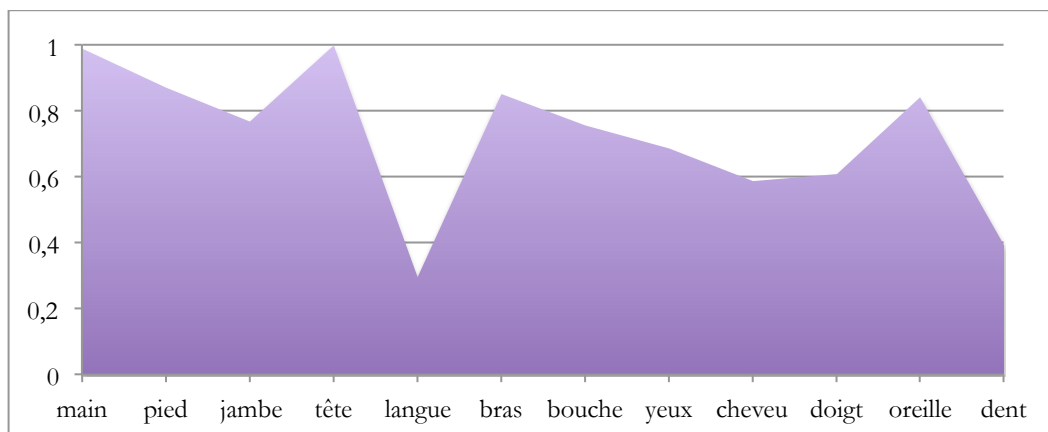


Gráfico 6.13. Muestra de compatibilidad léxica en francés. Modelo de reentrada leve. Informante 115. Centro de interés 01. *Parties du corps humain*

En síntesis, el modelo de reentrada leve se dibuja en las respuestas de los informantes que han aportado un número reducido de unidades léxicas. La mayoría corresponden a vocablos centrales que forman parte de la lista del léxico nuclear seleccionada a través del modelo de conjuntos difusos. Los procesos de reentradas son leves ya que el informante se desvía ligeramente del núcleo para, inmediatamente después, volver a entrar. Las vínculos entre las palabras son menos evidentes, pero cabe mencionar las relaciones de meronimia y espaciales. No obstante, en los tres idiomas se observan que las palabras no muestran una estructura organizada convencional: *boca-pie*, *barriga-mano* en español; *leg-heart* en inglés; *cheveu-doigt-oreille-dent* en francés.

6.4.2.2. Modelo de reentrada moderado

Este modelo combina procesos de reentrada leves y pronunciados. Se podría decir que los tres patrones localizados en este tipo de centros forman un continuo en el que se sitúan, por un lado, los patrones de reentrada leves y, por otro lado, los

pronunciados. Entre un tipo y otro se sitúan los patrones moderados. De esta forma, cuanto más rico sea el léxico de un hablante más, se orientará hacia un modelo pronunciado y viceversa, es decir, los informantes que aporten menos palabras presentarán, en general, procesos de reentradas más leves. Las entradas y salidas al núcleo son más numerosas con respecto al modelo anterior y se suceden vocablos con índices de compatibilidad inferiores y superiores a la media.

En el Gráfico 6.14 se ha seleccionado al informante número 39 por encontrarse en el umbral con respecto al listado de sujetos ordenados por número de palabras en el centro 1, concretamente en el puesto 85 de 171. El número de palabras aportadas es de 28, por lo que se sitúa en torno a la media en ese centro de interés (26,27 unidades léxicas por informante). Se observa que las seis primeras palabras del hablante son muy centrales y apenas se desvían del núcleo. En cambio, esa tendencia se rompe cuando el sujeto escribe la palabra *cerebro* con un índice de compatibilidad que baja al 0,209. A continuación, vuelve a entrar al núcleo de forma leve y desciende de nuevo con el vocablo *extremidad*, lexía que no se encuentra en el listado de aquellas más compatibles y cuenta con un índice de 0,149. Esos procesos de subidas pronunciadas se repiten en varias ocasiones acompañados de descensos y ascensos leves.

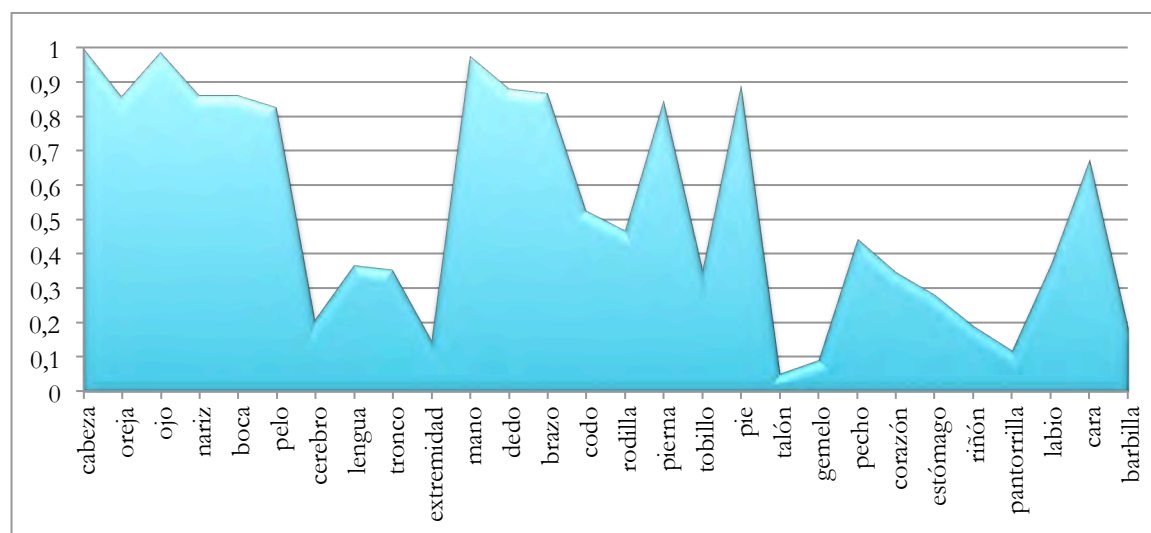


Gráfico 6.14. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 039. Centro de interés 01. *Partes del cuerpo humano*

Con respecto a la organización de los vocablos aportados, se constata que los ocho primeros hacen referencia a la cabeza y a partes que se sitúan en ella (*cabeza, oreja, ojo, nariz, boca, pelo, cerebro y lengua*), le siguen dos vocablos que actúan como

nexos (*tronco, extremidad*) que dan paso a las relaciones típicas de meronimia y holonimia (*mano-dedo, y brazo-codo* en el primer caso, y *dedo-brazo*, en el segundo). Como es frecuente en los listados, después de describir la extremidad del brazo, el informante describe la pierna (*rodilla, pierna, tobillo, pie, talón y gemelo*). A partir de este punto, el sujeto vuelve al núcleo con *pecho* y menciona tres órganos (*corazón, estómago y riñón*). Una vez que el informante ha descrito a grandes rasgos el cuerpo, aporta un vocablo relacionado con la pierna (*pantorrilla*) y comienza a describir más específicamente las partes de la cara (*labio y barbilla*).

En inglés, se ha incluido el ejemplo del sujeto 161 por presentar una media de palabras por centro de interés de 17,14, muy similar a la media del total de la muestra de estudio (14,64). En el Gráfico 6.15 se observan tres procesos de reentrada bruscos (*ankle-eye, leg-chin* y *ear-eyebrow*) combinados con otros leves y moderados (*eye-head-had-arm-leg*, entre otros). El informante es capaz de definir mejor los campos léxicos que en el modelo de reentrada leve, como muestran las palabras relacionadas con cabeza (*mouth, lip, face, teeth, nose, ear* y *eyebrow*). Además, el vocabulario es más rico ya que el sujeto aporta algún vocablo más descentralizado (*ankle* y *chin*) que no aparece incluido en los seis niveles de compatibilidad. Sin embargo, incluye otras secuencias cuya relación parece menos lógica (*leg-chin-finger-heart*).

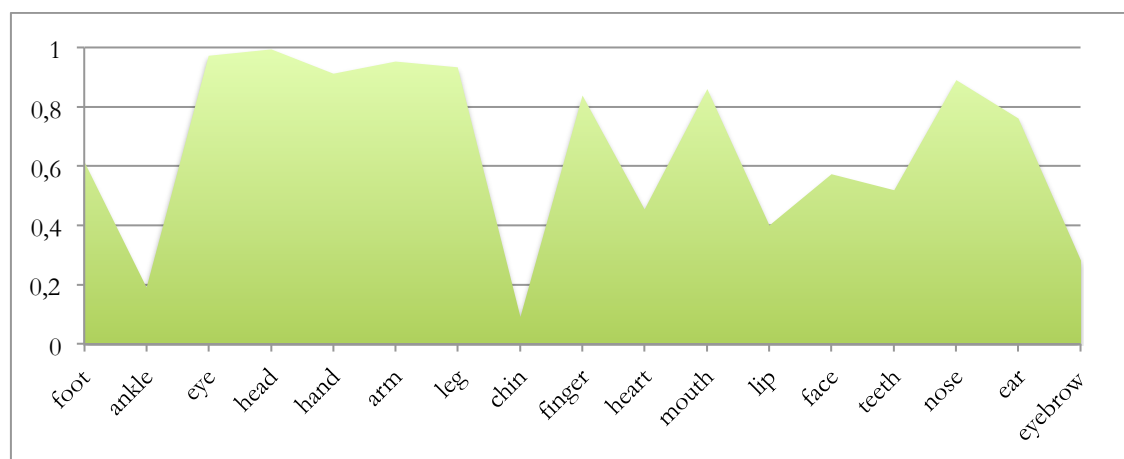


Gráfico 6.15. Muestra de compatibilidad léxica en inglés. Modelo de reentrada moderado. Informante 161. Centro de interés 01. *Human body parts*

En francés, las palabras del modelo pertenecen al informante 33 que ha escrito 20 palabras y el promedio en ese centro de interés es de 20,33 palabras por hablante. A simple vista, se observan dos reentradas abruptas compuestas por el conjunto de vocablos *cou-hanche-bras* y *épaule-épine dorsale*. En el primer caso, el

informante modifica la sucesión lógica de relaciones, es decir, pasa de escribir vocablos relacionados con la cabeza a mencionar *hanche* y luego *bras*. En el segundo supuesto, escribe una palabra nuclear como *épaule* seguida de una más específica *épine dorsale*.

En la misma línea que en los ejemplos anteriores, el informante intenta describir de forma ordenada las partes del cuerpo humano. Se detectan dos grupos que guardan una relación vertical (*bras-main-doigt-ongle* y *jambe-pied*) y un subgrupo en el que una palabra rompe esa relación vertical (*pied-nez-cheville*). En ese caso, parece que lo más lógico hubiera sido escribir *pied-cheville* y haber escrito *nez* en el subgrupo relacionado con *cabeza*. El resto de vocablos está relacionado por pares *épaule-épine dorsale*, *cul-poitrine*.

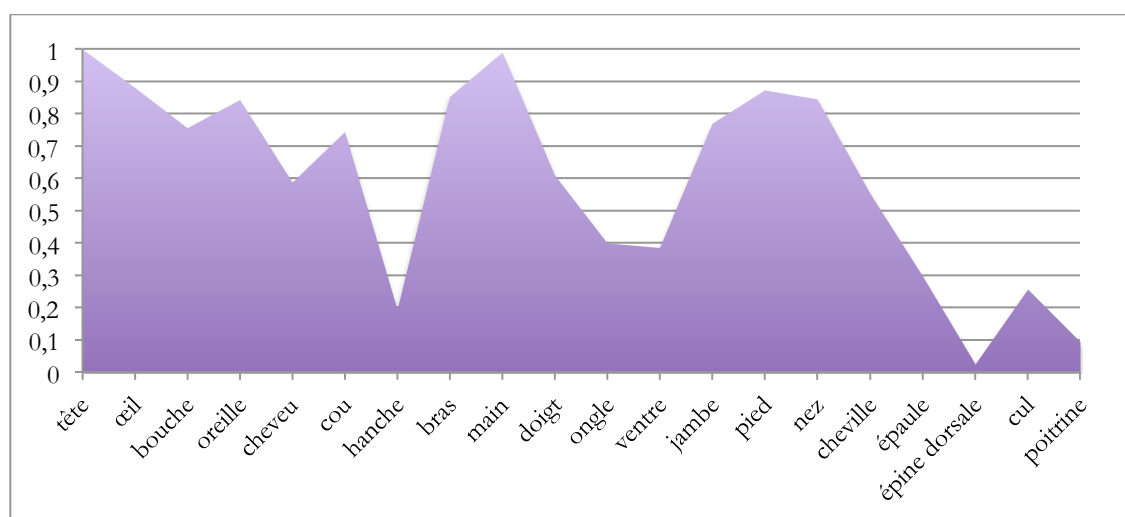


Gráfico 6.16. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 033. Centro de interés 01. *Parties du corps humain*

6.4.2.3. Modelo de reentradas pronunciado

Este modelo se presenta como el opuesto al modelo de reentrada leve. Los listados suelen ser bastante productivos y contar con palabras específicas, especializadas o excepcionalmente provenientes de asociaciones individuales. Los procesos de reentrada son numerosos e incluyen palabras con un alto índice de compatibilidad intercaladas con vocablos por debajo de la media de compatibilidad del centro en cuestión. La muestra elegida corresponde al informante 109, sujeto que cuenta con el mayor número de palabras aportadas en el centro 1.

A partir del vocablo 7, las palabras se alejan del núcleo hasta el vocablo 11, *diente canino* con un índice muy inferior a la media, concretamente del 0,007. La reentrada se produce con una palabra tan compatible como *oreja* con un índice de 0,858. Esos procesos de reentrada son muy frecuentes debido a que el informante aporta un listado con palabras genéricas que completa con palabras técnicas: *estómago* relacionada *páncreas* y *bilis*, *rodilla* y *espinilla*²⁸⁹ conectadas con *tibia* y *peroné*.

Por tanto, se trata de hablantes con un ICoDD muy elevado que entran y salen del núcleo constantemente. Se asocia con el vocabulario especializado ya que normalmente el hablante menciona una palabra general y luego especifica con una palabra más particular y menos centralizada. En esta ocasión son muy frecuentes las relaciones de meronimia: *mano-dedo*, *boda-diente*, *cabeza-pelo*, etc.

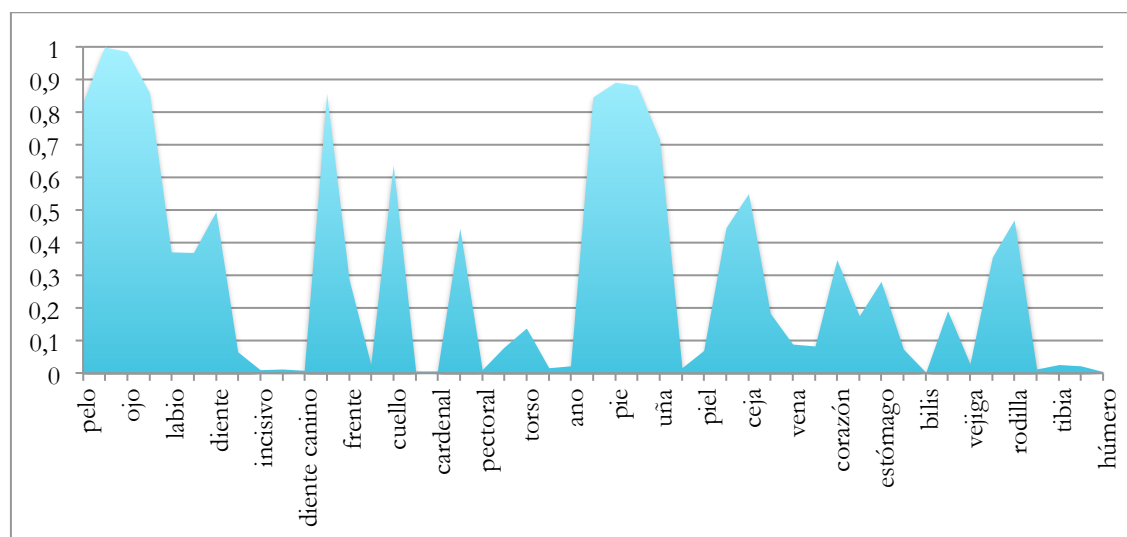


Gráfico 6.17. Muestra de compatibilidad léxica. Modelo de reentrada pronunciado. Informante 109. Centro de interés 01. *Partes del cuerpo humano*

Para poder consultar las relaciones entre las palabras aportadas por el sujeto 109, se adjunta a continuación la Tabla 6.10 con la posición y el nivel de compatibilidad²⁹⁰ de cada vocablo. Se observa que la secuencia de unidades léxicas presenta una organización más clara en este caso que la que se observaba en los informantes que han aportado pocas palabras. De esta forma, se observan subcampos dentro de esa lista como por ejemplo: *vena-arteria-corazón*; luego continúa

²⁸⁹ Este vocablo es considerado por el DRAE como familiar pero entra en lo que se considera léxico específico ya que se refiere a una parte concreta de la pierna.

²⁹⁰ El informante ha escrito un total de 47 palabras lo cual imposibilita que puedan consultar de forma clara todas las palabras en un diagrama.

con órganos internos *hígado*, *estómago*, *páncreas*, vuelve con un merónimo de hígado, *la bilis*, y continúa con otros dos órganos: *riñón* y *vejiga*.

A continuación, vuelve a entrar en el núcleo con la palabra *tobillo* que es la articulación que une el pie y la pierna, sigue con otra articulación, *la rodilla*, después cita la parte anterior de la pierna, *la espinilla*, vuelve con un hueso que forma parte del tobillo y la rodilla, *la tibia*, el cual asocia con otro de los huesos del tobillo, *el peroné* y, por último, acaba con un hueso del brazo, el *húmero*. Esa es la última palabra del listado pero quizás con más tiempo el informante podría haber enumerado los huesos y partes del brazo.

Posición	Vocablo	Compatibilidad	Posición	Vocablo	Compatibilidad
1	pelo	0,82571	25	pie	0,89091
2	cabeza	0,998292	26	dedo	0,879814
3	ojo	0,984674	27	uña	0,714797
4	nariz	0,859977	28	arruga	0,014649
5	labio	0,370333	29	piel	0,067378
6	lengua	0,367802	30	pestaña	0,444669
7	diente	0,493833	31	ceja	0,548354
8	muela	0,06411	32	barbilla	0,182193
9	incisivo	0,00868	33	vena	0,087388
10	paleta	0,011892	34	arteria	0,081539
11	diente canino	0,007102	35	corazón	0,346171
12	oreja	0,857713	36	hígado	0,174066
13	frente	0,285516	37	estómago	0,28061
14	sien	0,024803	38	páncreas	0,072189
15	cuello	0,63746	39	bilis	0,002003
16	lesión	0,004882	40	riñón	0,190164
17	cardenal	0,004595	41	vejiga	0,027686
18	pecho	0,442559	42	tobillo	0,354091
19	pectoral	0,009298	43	rodilla	0,468422
20	abdomen	0,078201	44	espinilla	0,010275
21	torso	0,136837	45	tibia	0,024726
22	pubis	0,014982	46	peroné	0,021121
23	ano	0,021317	47	húmero	0,00372
24	pierna	0,845749			

Tabla 6.10. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica. Informante 109. Centro de interés 01. *Partes del cuerpo humano*

En inglés, los informantes 145 y 24 son aquellos que han aportado más palabras en este centro de interés, concretamente 32 lexías. Al igual que en el ejemplo anterior, se distinguen diferentes subcampos léxicos que corresponden a estructuras similares a las encontradas en español: articulaciones y huesos del brazo (*elbow*, *wrist*,

hand, finger, waist) seguidas de aquellas relacionadas con la pierna (*knee, ankle, calf muscle, foot, feet, toe*) y con órganos internos (*liver, heart, lung, kidney, stomach*).

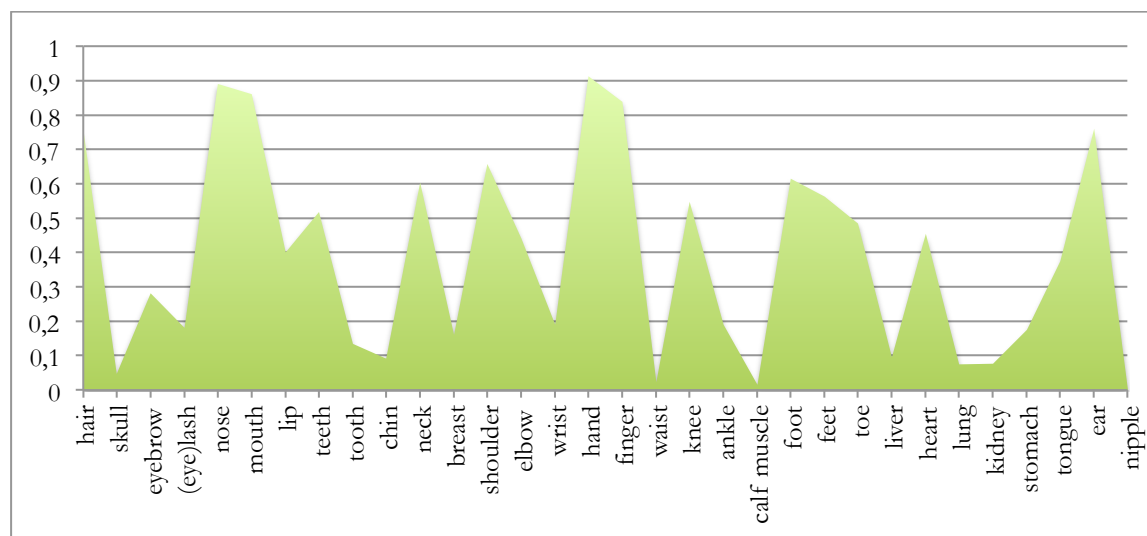


Gráfico 6.18. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado. Informante 146.
Centro de interés 01. *Human body parts*

En francés, se presentan las 25 palabras aportadas por el informante número 28. El informante comienza con una palabra poco usual, *cheville*, continúa con las partes del cuerpo relacionadas con la cabeza (*cou, visage, yeux, oreille, bouche, nez, cil* y *sourcil*), desciende al brazo, que acompaña con dos palabras que guardan una relación de meronimia (*main* y *doigt*) y unas palabras más tarde vuelve a repetir este mismo proceso pero con la pierna (*genou* y *pied*). Por último, hace referencia a dos órganos (*poumon* y *estomac*) y termina reentrando de nuevo al núcleo con el vocablo *cheveu*.

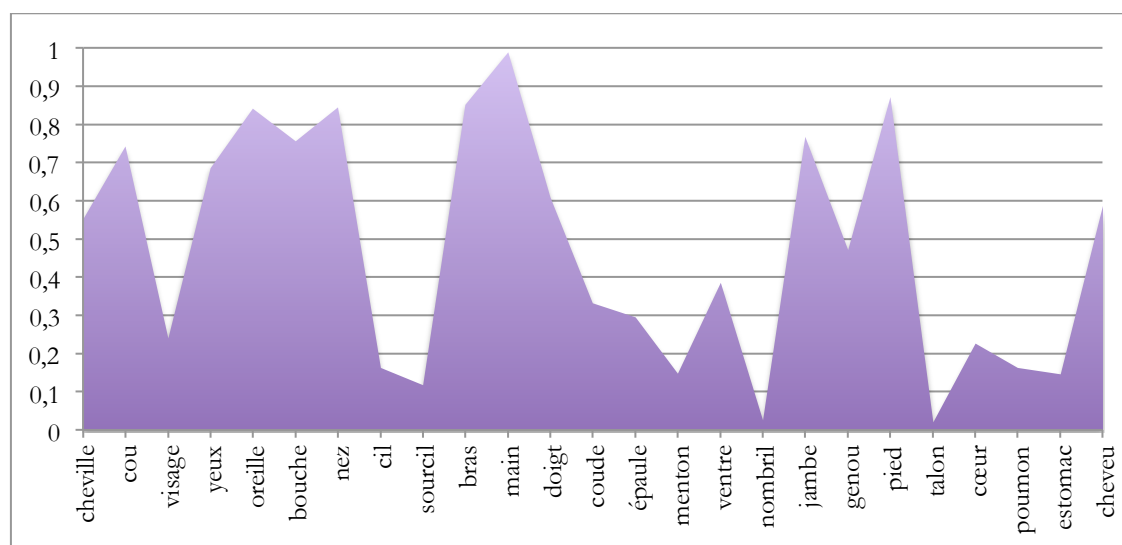


Gráfico 6.19. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado. Informante 028.
Centro de interés 01. *Parties du corps humain*

6.4.2.4. Comparación entre la secuencia léxica de un informante en español y en lengua extranjera

En este apartado se presenta un análisis comparativo entre las palabras aportadas por un mismo informante en español y en lengua extranjera. Se ha seleccionado a un hablante que llama la atención por la gran cantidad de palabras aportadas en francés, un total de 49, que supera al sujeto que más unidades léxicas ha escrito en español, 47. El objetivo es comprobar si en español y en lengua extranjera ha realizado asociaciones léxicas similares para confirmar la hipótesis de este apartado o si, por el contrario, se ha limitado a aportar palabras que conoce en francés.

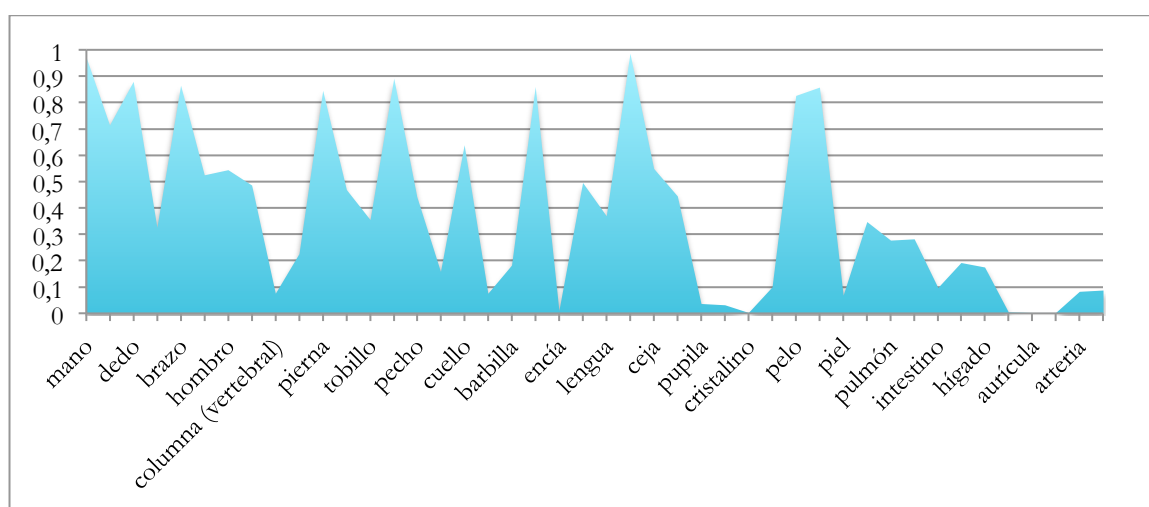


Gráfico 6.20. Muestra de compatibilidad léxica en español. Caso extremo. Informante 038. Centro de interés 01. *Partes del cuerpo humano*

A simple vista, se observa que en español y francés los procesos de reentrada al núcleo son bruscos, lo que muestra que presentan un vocabulario descentralizado que complementa al léxico nuclear. En el Gráfico 6.19 y el Gráfico 6.20 se observa, al final, una tendencia de reentrada algo más moderada ya que la palabra que vuelve al núcleo supera la media de compatibilidad levemente. Se podría decir que en este tipo de centros el informante nunca llega a un punto de no retorno sino que siempre vuelve al centro para luego aportar palabras más específicas.

En la Tabla 6.11 se han marcado los vocablos que aparecen en las listas en español pero que no se encuentran en el listado en francés. El número total de vocablos diferentes es de 15, lo que supone un 34,09 % del total. Por tanto, la mayoría de vocablos (65,91 %) coincide en ambos listados sobre todo en las primeras

posiciones. Así, entre las diez primeras vocablos más compatibles solo aparece uno en español que no está en francés, *muñeca*.

Posición	Vocablo	Compatibilidad	Posición	Vocablo	Compatibilidad
1	mano	0,974241	23	lengua	0,367802
2	uña	0,714797	24	ojo	0,984674
3	dedo	0,879814	25	ceja	0,548354
4	muñeca	0,323794	26	pestaña	0,444669
5	brazo	0,865433	27	pupila	0,035521
6	codo	0,524893	28	iris	0,030091
7	hombro	0,543855	29	crystalino	0,002693
8	espalda	0,485403	30	cabello	0,100015
9	columna (vertebral)	0,073678	31	pelo	0,82571
10	culo ²⁹¹	0,224242	32	oreja	0,857713
11	pierna	0,845749	33	piel	0,067378
12	rodilla	0,468422	34	corazón	0,346171
13	tobillo	0,354091	35	pulmón	0,275321
14	pie	0,89091	36	estómago	0,28061
15	pecho	0,442559	37	intestino	0,096337
16	ombligo	0,158295	38	riñón	0,190164
17	cuello	0,63746	39	hígado	0,174066
18	garganta	0,075404	40	bronquio	0,005201
19	barbilla	0,182193	41	aurícula	0,001905
20	boca	0,859675	42	ventrículo	0,00186
21	encía	0,010206	43	arteria	0,081539
22	diente	0,493833	44	vena	0,087388

Tabla 6.11. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica en español. Informante 038. Centro de interés 01. *Partes del cuerpo humano*

En cuanto a la organización de las palabras, se observa cierta similitud en los dos listados ya que existen ciertos grupos de palabras que siguen un orden casi idéntico. Por ejemplo, en español el informante ha escrito *mano-uña-dedo-muñeca-brazo* y en francés *main-bras-doigt-ongle*. Otra secuencia léxica similar está formada por vocablos que hacen referencia al cuello y la boca: *cuello-garganta-barbilla*²⁹²-*boca-encía-diente-lengua* y *cou-gorge-bouche-lèvre-gencive-dent-langue*.

²⁹¹ El vocablo *culo* se corresponde en francés *cul* y *fesse* aunque este último también hace referencia a *nalga*.

²⁹² Aparece en posiciones posteriores en francés, concretamente en la posición 31, tras el vocablo *joue*.

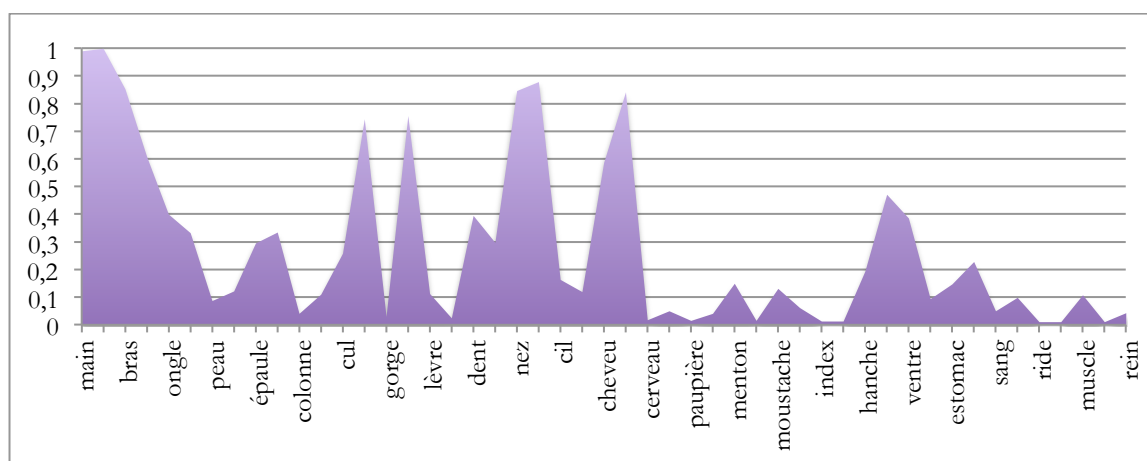


Gráfico 6.21. Muestra de compatibilidad léxica en francés. Caso extremo. Informante 038. Centro de interés 01. *Parties du corps humain*

En francés, se han destacado 19 palabras que no aparecen en el listado en español, lo que equivale a un 38,78 % del total de palabras aportadas por el informante. Nuevamente la mayoría de las palabras, un 61,22 %, coincide en ambos listados. En este idioma el informante ha aportado más información con respecto a los nombres de los dedos (*pouce*, *index* y *petit doigt*), vocablos muy compatibles que no aparecen en español (*tête*, *lèvre* y *nez*) y otros menos compatibles como (*cerveau*, *crâne*, *paupière*, *ride*, *nerf*, *muscle*, *sang* y *hanche*) que no forman parte del léxico nuclear.

Posición	Vocablo	Compatibilidad	Posición	Vocablo	Compatibilidad
1	main	0,988913	26	oreille	0,84104
2	tête	0,998753	27	cerveau	0,016203
3	bras	0,85177	28	crâne	0,048753
4	doigt	0,60799	29	paupière	0,015086
5	ongle	0,398747	30	joue	0,038534
6	coude	0,331308	31	menton	0,147554
7	peau	0,086626	32	barbe	0,013671
8	poil	0,121328	33	moustache	0,129676
9	épaule	0,294459	34	pouce	0,060853
10	dos	0,332973	35	index	0,012499
11	colonne	0,039772	36	petit doigt	0,012152
12	os	0,109179	37	hanche	0,191319
13	cul	0,256465	38	genou	0,470167
14	cou	0,742817	39	ventre	0,384844
15	gorge	0,029166	40	poitrine	0,09244
16	bouche	0,755076	41	estomac	0,146545
17	lèvre	0,11232	42	cœur	0,226712
18	gencive	0,024305	43	sang	0,049889
19	dent	0,393229	44	fesse	0,096573
20	langue	0,296369	45	ride	0,009722

21	nez	0,844553	46	nerf	0,00951
22	œil	0,878726	47	muscle	0,10609
23	cil	0,162323	48	foie	0,009114
24	sourcil	0,118172	49	rein	0,042281
25	cheveu	0,586952			

Tabla 6.12. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica en francés. Informante 038. Centro de interés 01. *Parties du corps humain*

Para concluir, resulta llamativo que en francés el informante 38 haya escrito cinco palabras más que en su lengua materna. Este comportamiento se podría justificar por el nivel de competencia lingüística del informante y por las pautas seguidas en la realización de la prueba. Por un lado, se trata de un hablante bilingüe licenciado en Traducción e Interpretación que cursa la especialidad del Máster en Profesorado de Lengua Extranjera (Francés). Por otro lado, la propia metodología de la recogida de los datos hace reflexionar al hablante primero en lengua materna y luego en lengua extranjera lo que puede condicionar en cierto modo las respuestas y provocar que se repitan las estructuras léxicas.

El estudiante vuelve a escribir las palabras en el mismo centro de interés pero en otra lengua tras contestar a nueve listados diferentes, lo que equivale a un mínimo de 18 minutos más tarde. Si el informante no contase con un repertorio lingüístico abundante en lengua extranjera se limitaría a aportar algunas palabras que suelen coincidir con las más compatibles y, por tanto, presentaría un modelo de reentrada leve. En este caso, el informante posee un amplio bagaje lingüístico en francés lo que le permite volver a su memoria a corto plazo para producir un léxico similar al español y ampliar el número de respuestas con cinco palabras más.

6.4.3. Modelo de reentrada de los centros de interés relacionales

Los elementos que componen los modelos relacionales no se encuentran tan bien definidos y no se identifican de forma tan clara con el prototipo como los anteriores, es decir, los inclusivos. Al igual que Hernández Muñoz (2006: 151) propone cambiar la proposición lógica a «x es parte de Y» a la que corresponden los centros inclusivos, por ejemplo «la pizarra es parte de la escuela» por «x está relacionado con Y» o «x tiene algún tipo de relación con Y». De esta forma, el

desayuno podría estar relacionado con el centro *La escuela* ya que se realiza durante el recreo.

En este caso, los centros no corresponden escrupulosamente a la clasificación de centros relacionales de Hernández Muñoz (2006) pero comparten la ecuación anterior. Estos centros se caracterizan por un repertorio lingüístico heterogéneo que lleva a un léxico nuclear más amplio y a unas relaciones más dispersas. Los informantes encuentran mayor dificultad a la hora de identificar el prototipo e introducen asociaciones indirectas y personales que llevan a diversificar el vocabulario.

Siguiendo la mecánica anterior, se han analizado los informantes situados en diferentes niveles en cuanto a número de palabras aportadas. Sin embargo, solo se han definido dos tipos de modelos de reentrada: el modelo de reentrada moderado en los casos en los que los informantes aportan pocas palabras y el modelo de reentrada pronunciado con desviación lineal en el caso contrario. Entre estos dos extremos se sitúa el resto de informantes que presenta un modelo intermedio. El centro tomado como referencia para mostrar los resultados es *La escuela: muebles y materiales* por estar directamente relacionado con la práctica docente de los estudiantes del Máster en Profesorado y por ser un centro que se encuentra en una posición central de la escala de centros según su prototipicidad.

6.4.3.1. Modelo de reentrada moderado

En los centros relacionales, la variedad léxica es mayor y el vocabulario nuclear está menos definido ya que los informantes aportan un número similar de palabras pero un número mayor de vocablos. Por esta razón, los informantes con un reducido caudal léxico no se ciñen únicamente a aquellos términos nucleares sino que se desvían hacia asociaciones indirectas en seguida. El hablante no tiene la misma percepción que en los centros inclusivos de que está cruzando la línea de incertidumbre por lo que es frecuente encontrar más asociaciones individuales.

En español se presentan los datos del informante 60 que ha escrito tan solo 13 palabras en el centro *La escuela: muebles y materiales*, mientras que el promedio para ese centro es de 21,74 unidades léxicas por sujeto. A grandes rasgos, todo el vocabulario

aportado por el sujeto corresponde al léxico nuclear a excepción de la palabra *mapa*. La secuencia léxica no está organizada y dividida en subcampos a excepción de dos pares de palabras *alumno-profesor* y *estuche-mochila*. Por ejemplo, la primera palabra que aporta el informante es *pizarra* pero no menciona la palabra *tiza* hasta la duodécima posición. En cambio, como se mostrará en los tipos de relaciones del programa DispoGrafo, *pizarra* y *tiza* mantienen una fuerte relación.

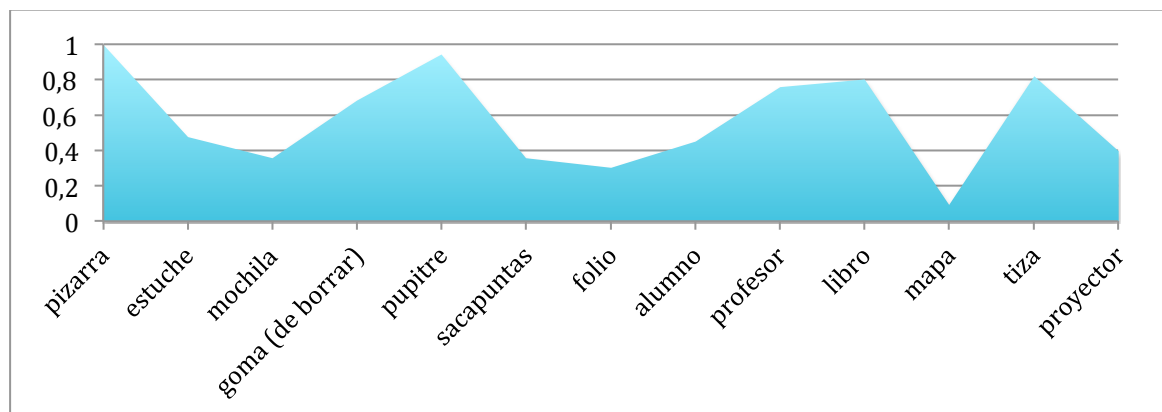


Gráfico 6.22. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 060. Centro de interés 04. *La escuela: muebles y materiales*

En inglés, el ejemplo muestra las respuestas del informante 118 que ha aportado 8 palabras mientras que el promedio en este centro de interés es de 14,13 unidades léxicas. Al igual que en el modelo en español, todos los vocablos a excepción de uno, *bin*, pertenecen al léxico considerado nuclear. Las tres primeras palabras formarían parte de las herramientas que tiene un informante para escribir y borrar (*pencil*, *pen* y *rubber*) aunque el orden lógico hubiera sido (*pen*, *pencil* y *rubber*) ya que este último sirve para borrar la escritura del lápiz. A continuación le sigue parte del mobiliario escolar (*table*, *chair*, *door*, *window* y *bin*).

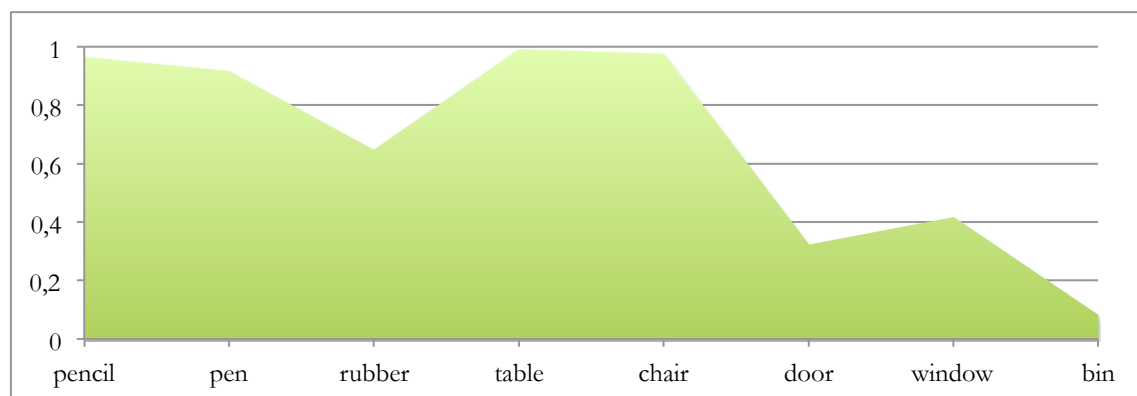


Gráfico 6.23. Muestra de compatibilidad léxica en inglés. Modelo de reentrada moderado. Informante 118. Centro de interés 04. *The school: furniture and school material*

En francés, el informante 170 junto con el 115 son los que menos palabras han aportado, un total de 9, mientras que el promedio para ese centro de 15,38. Las unidades léxicas pertenecen en su mayoría al léxico categorizado como nuclear a excepción de *DVD* y *classeur*. La organización atiende, en primer lugar, al mobiliario escolar (*table*, *tableau*, *chaise*, *televisión* y *DVD*) y, a continuación, al material escolar (*livre*, *classeur*, *dossier* y *cahier*). Al igual que en el ejemplo en lengua inglesa, las dos primeras palabras comparten la misma raíz (*table* y *tableau*).

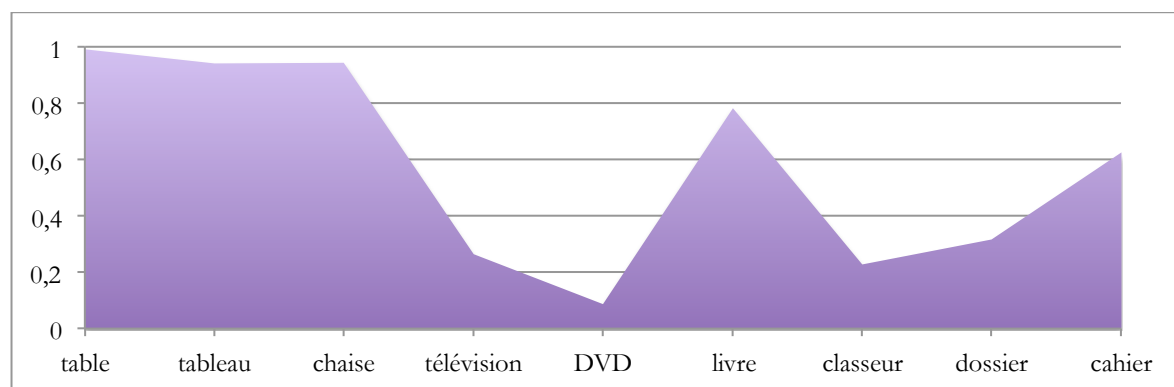


Gráfico 6.24. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 170. Centro de interés 04. *L'école: meubles et matériel scolaire*

En los tres ejemplos se ha puesto de manifiesto que en los centros de interés relacionales los informantes que han escrito pocas palabras no se limitan a mencionar vocablos nucleares con un alto índice de compatibilidad que lleva a procesos de reentrada leve como ocurría en los centros inclusivos. En este caso, en las tres lenguas se incluye al menos una palabra no nuclear y los procesos de reentrada son más acentuados. Ello viene a corroborar que la estructura básica de este centro no está tan bien definida como en los centros inclusivos.

6.4.3.2. Modelo de reentrada pronunciado con desviación lineal

Para la muestra de este modelo se han seleccionado a los informantes con mayor caudal léxico. Los resultados son similares a los obtenidos en los centros inclusivos salvo un comportamiento que se repite en todos los hablantes analizados: a medida que se van escribiendo palabras los procesos de reentrada se alejan de forma progresiva del núcleo. Es decir, aunque la reentrada al núcleo sea fuerte poco a poco irá menguando y debilitando su compatibilidad.

En español, el informante 89 ocupa el segundo puesto en productividad con 35 vocablos (el primero corresponde al informante 34 que ha escrito 36 palabras). Entre las diez primeras palabras hay siete que se refieren a material y mobiliario escolar que el sujeto ha relacionado con el usuario (*pupitre-alumno-compañero* y *profesor-tiza*). A continuación, le sigue un subcampo referido al material escolar (*ordenador, lápiz, papel, libreta, hoja, regla, cartabón, compás, estuche, boli(grafo), rotulador, pincel* y *libro*), el accesorio donde guarda el material (*mochila*) y la indumentaria (*uniforme*). Este último lo asocia con los *deberes* y, seguidamente por contraposición, con *recreo*. Por último, menciona partes del colegio (*biblioteca* y *comedor*) y tras una asociación individual (*babero*) pasa a enumerar de nuevo el material escolar (*plastilina, cartulina* y *goma de borrar*).

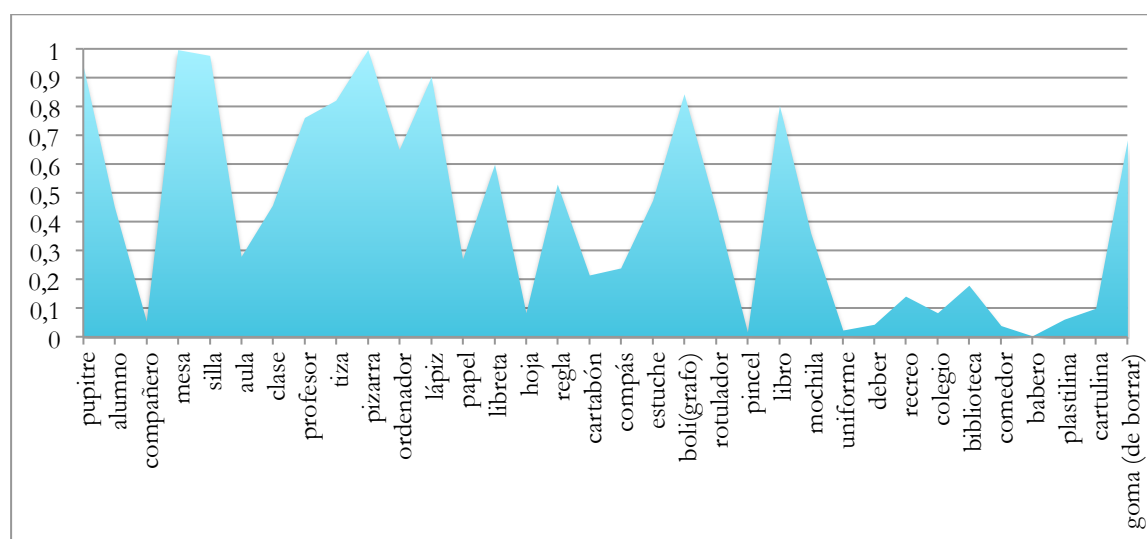


Gráfico 6.25. Muestra de compatibilidad léxica en español. Modelo de reentrada pronunciado con desviación lineal. Informante 060. Centro de interés 04. *La escuela: muebles y materiales*

En inglés, el informante que más palabras ha aportado ha sido el 146 con 29 palabras. Los procesos de reentrada iniciales se producen porque el sujeto hace referencia a una parte determinada de la escuela, *gym*, a un puesto de trabajo específico, *headmaster*, y a un material escolar concreto, *sheet of paper* (lo más usual es escribir el genérico *notebook*).

En cambio, el hablante escribe una breve secuencia de palabras con un bajo índice de compatibilidad sin reentrar al núcleo cuando habla de actividades concretas de la escuela: *essay, assessment, deadline, oral examination* y *written examination*. Vuelve al núcleo con la palabra *computer* y se desvía al nombrar los tipos de asignaturas (*maths, laboratoy, physics, subject* y *language*). Se observa la tendencia del hablante a escribir pares

de palabras que comparten la misma raíz (*pencil-pen* y *notebook-book*) pero también se observa la capacidad que tiene para aportar unidades léxicas relacionadas (*blackboard-chalk* y *schoolbag-rucksack*).

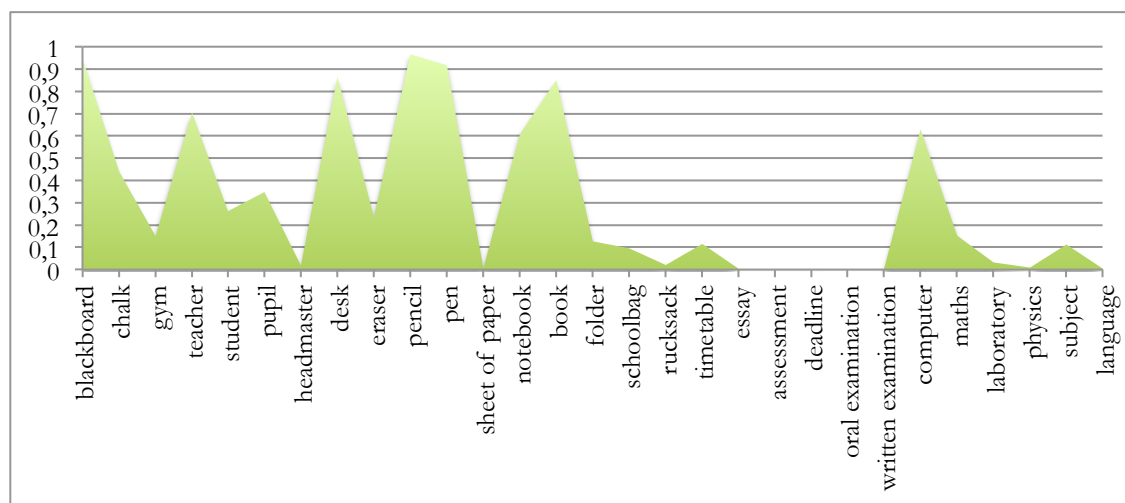


Gráfico 6.26. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado con desviación lineal. Informante 146. Centro de interés 04. *The school: furniture and school material*

En francés, el ejemplo se realiza a partir del informante que ocupa el segundo puesto en productividad con 22 palabras. En esta ocasión, el hablante comienza a aportar unidades léxicas relacionadas con los muebles (*table*, *chaise*, *fautuil* y *tableau*) sigue con el material escolar (*crayon*, *effaceur* y *stylo*) y vuelve al subcampo léxico del mobiliario mezclándolo con otros vocablos. Las tres primeras desviaciones del núcleo se producen porque menciona un tipo de asiento específico, *fautuil*, una palabra nuclear menos compatible, *effaceur*, y un vocablo que hace referencia al edificio propiamente dicho, *mur*.

En este modelo llama la atención que, a partir de un determinado momento, el informante se aleja del núcleo y solo realiza reentradas muy vagas. Ello se produce porque el sujeto es capaz de proporcionar una descripción más específica del edificio (*toit*, *plafond* y *plancher*) y porque aporta palabras cuya relación con el prototipo es indirecta y apenas es compartida por el resto de informantes (*clé*, *lumière*, *ballon* y *balle*). No obstante, si el sujeto hubiese continuado la lista seguramente volvería al núcleo con algún vocablo con menor índice de compatibilidad que los primeros.

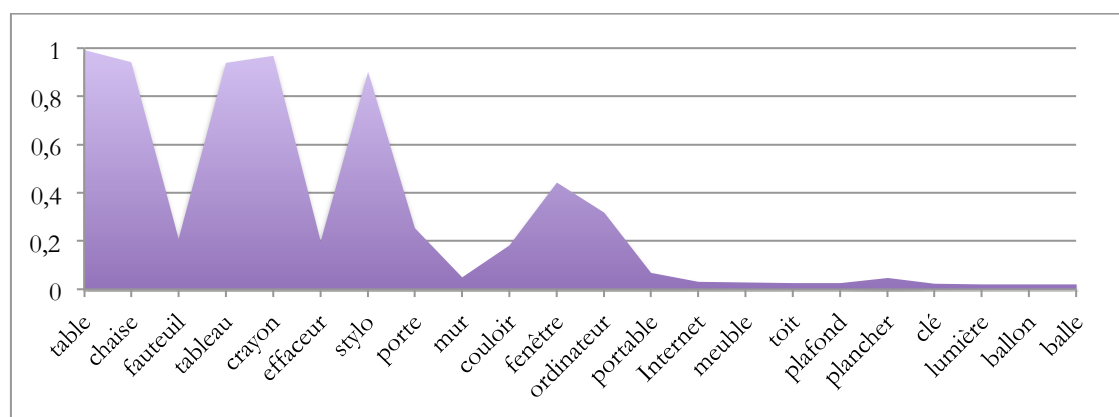


Gráfico 6.27. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado con desviación lineal. Informante 034. Centro de interés 04. *L'école: meubles et matériel scolaire*

En esta ocasión, se ha optado por elegir el modelo del informante que ocupa el segundo puesto en el rango de productividad y no el primero porque la muestra de compatibilidad presenta un efecto de descentralización que se prolonga más que en el caso del español y el inglés. El informante 38 ha aportado un total de 30 palabras y dibuja un gráfico muy similar al de las otras lenguas. En cambio, la idea es transmitir que a grandes rasgos los sujetos que han escrito muchas palabras siguen el modelo de reentrada pronunciado con una desviación lineal pero hay casos en los que realizan una serie de asociaciones individuales que les hacen perder la referencia con respecto al núcleo.

6.4.4. Ejemplo de clasificación del centro de interés *Ordenadores e internet* según los procesos de reentrada

En este centro de interés los procesos de reentrada siguen los patrones de los centros relacionales: el modelo de reentrada moderado y modelo de reentrada pronunciado con desviación lineal. Comparte con este tipo de centros características como el bajo índice de cohesión del centro y la heterogeneidad de las respuestas que conlleva a un vocabulario nuclear más amplio. Asimismo, se sitúa en el extremo de la clasificación realizada según la prototipicidad de los centros por lo que resulta interesante comparar el comportamiento con el centro que se encuentra en el otro extremo, *Partes del cuerpo humano*, y con un centro con unos resultados intermedios, *La escuela: muebles y materiales*.

6.4.4.1. Modelo de reentrada moderado

La organización léxica de los informantes con una producción léxica baja presenta unas desviaciones desde y hacia el núcleo moderadas. El hablante número 60 representado en el Gráfico 6.28 ha aportado tan solo 10 palabras mientras que el promedio por la muestra en ese centro de interés es de 19,24 unidades léxicas por sujeto. Todas las palabras son consideradas nucleares ya que es el segundo centro en número de unidades léxicas seleccionadas con un total de 69, situándose por detrás del octavo centro, *Profesiones y oficios*, con 70 palabras.

El informante comienza aportando datos sobre diferentes dispositivos del ordenador: el dispositivo de salida (*monitor*), dispositivo apuntador (*ratón*) y dispositivos periféricos (*impresora* y *escáner*). Tras un paréntesis de dos palabras (*archivo* y *portátil*) cita un dispositivo de almacenamiento (*disco duro*) menciona el término *página web* y vuelve con un dispositivo de entrada (*teclado*) y otro dispositivo de almacenamiento (*pendrive*). A grandes rasgos, el informante ha mencionado en siete vocablos algunos de los dispositivos más frecuentes de un ordenador.

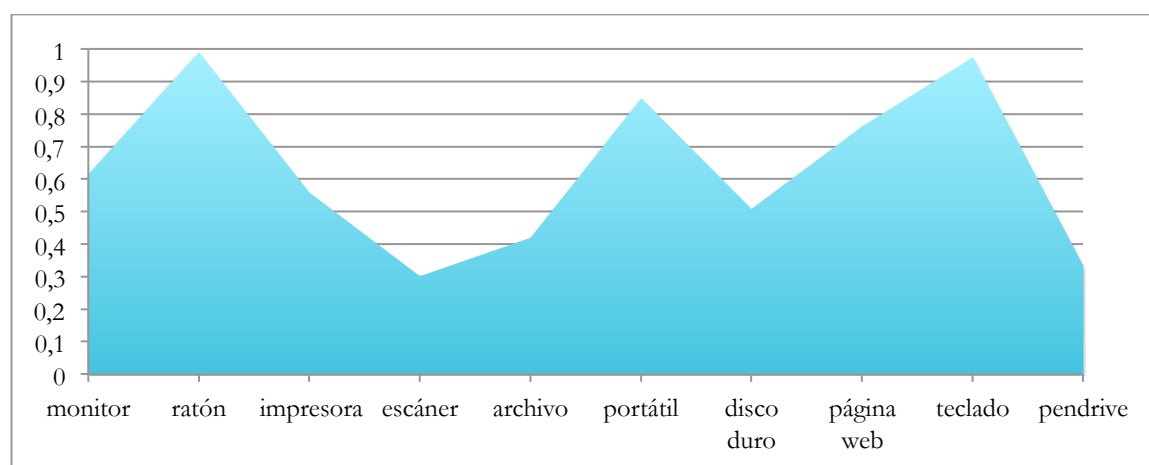


Gráfico 6.28. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 060. Centro de interés 09. *Ordenadores e internet*

En inglés, se muestran las ocho respuestas del informante 119. El promedio de palabras en este centro por la muestra general es de 15,01 y es el que presenta mayor número de vocablos nucleares (54) junto con el quinto centro, *The city*. Entre las palabras mencionadas se destacan dos que no pertenecen a ese listado de palabras compatibles (*charger* y *notebook*) lo que provoca que el proceso de reentrada sea mayor. La secuencia de palabras es similar a la presentada en español: el informante

comienza con *website*, a continuación menciona dos tipos de ordenador (*personal computer* y *laptop*), sigue con cuatro dispositivos (*charger*, *mouse*, *keyboard* y *screen*) y finaliza con otro tipo de ordenador (*notebook*).

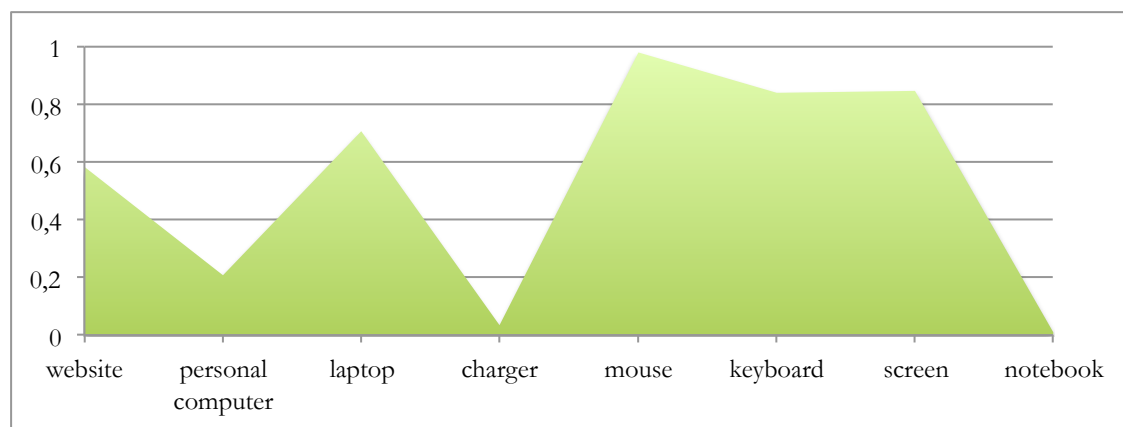


Gráfico 6.29. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 119. Centro de interés 09. *Computer and Internet*

En francés, se ha seleccionado a uno de los informantes que menos palabras ha aportado, concretamente el sujeto número 70 que ha escrito 9 palabras. En este centro el promedio de respuestas por el total de la muestra es de 15,05 y ocupa el tercer puesto en cuanto a número de palabras nucleares (28) junto con el octavo, *Professions et métiers*. No se incluyen dentro de la lista de vocabulario nuclear las unidades léxicas siguientes: *programme*, *courrier électronique* y *scanner*. Como ocurre en los casos anteriores, el hablante aporta datos sobre dispositivos pero no de forma ordenada ya que *disquette* aparece en primer lugar y de forma intercalada aparece *CD* y, por último *imprimante* y *scanner*.

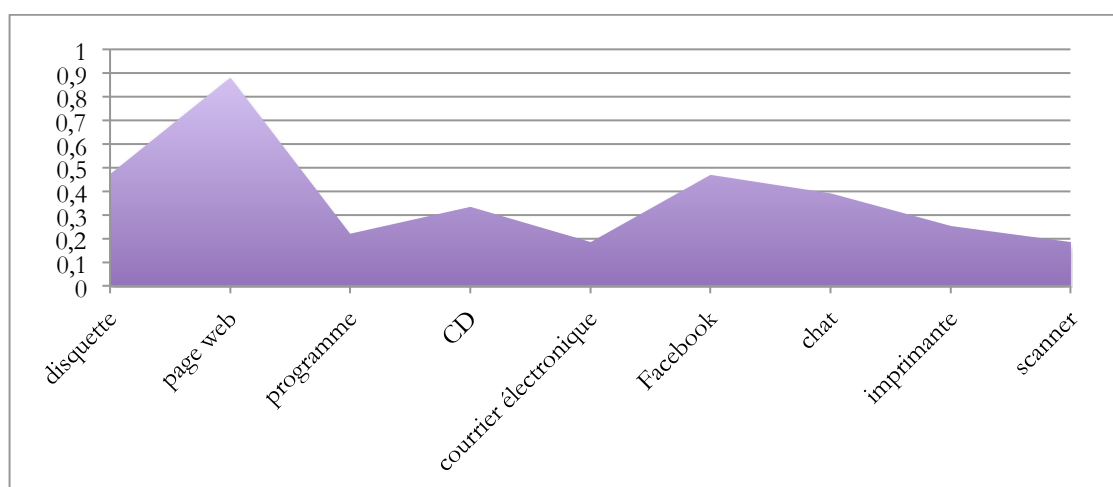


Gráfico 6.30. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 070. Centro de interés 09. *Ordinateurs et Internet*

En este centro de interés es singular la alta compatibilidad de la red social *Facebook* ya que muestra el uso de internet y los ordenadores entre los jóvenes. En este sentido, es imprescindible no eliminar esas palabras en el proceso de codificación de los datos para analizar cómo organizan los hablantes las palabras para representar la realidad. La riqueza léxica encontrada en los tres idiomas, aunque en menor medida en francés, provoca una mayor diversidad en las palabras nucleares y, por ende, que los procesos de reentrada sean más acentuados.

6.4.4.2. Proceso de reentrada pronunciado con desviación lineal

Los informantes con mayor productividad léxica realizan procesos de reentrada más acentuados. No obstante, al igual que ocurre con el modelo presentado de los centros relacionales, a medida que el hablante aporta palabras pierden compatibilidad por lo que la reentrada no es tan acentuada. Además, es frecuente que el sujeto escriba una secuencia de unidades léxicas con un índice de compatibilidad bajo para describir algún aspecto concreto o para aportar palabras fruto de asociaciones individuales.

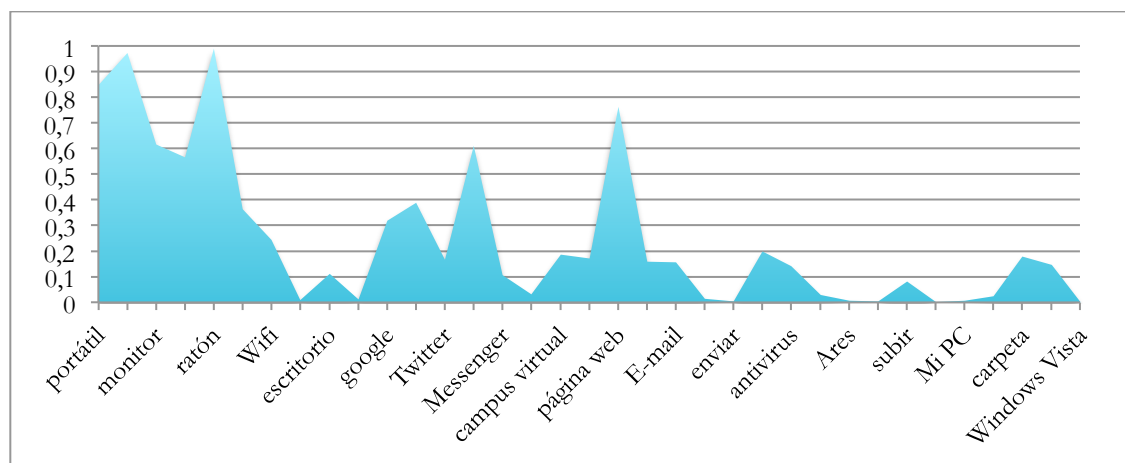


Gráfico 6.31. Muestra de compatibilidad léxica en español. Modelo de reentrada pronunciado con desviación lineal. Informante 016. Centro de interés 09. *Ordenadores e internet*

En la Tabla 6.13 se observa que los subcampos léxicos creados por el informante número 16 están mejor definidos que en aquellos que muestran un caudal léxico más pobre. El sujeto comienza mencionando un tipo de ordenador, *portátil*, para seguir con nombres de dispositivos. Se muestran subcampos léxicos relacionados con las formas de acceso a internet (*cable*, *Wifi* y *conexión telefónica*) como

preludio a los nombres de páginas web (incluido el genérico) en los que se incluyen: *Google, Tuenti, Twitter, Facebook y Messenger*. En ocasiones parece que el informante estuviera describiendo una acción que realiza: *página web, contraseña, e-mail, bandeja de entrada, enviar, virus, antivirus, Emule, Ares, descarga directa*, etc. Tras escribir *antivirus* el informante localiza dos programas de intercambios de archivos donde se pueden descargar programas de todo tipo, incluidos programas de antivirus (*Emule y Ares*).

Posición	Vocablo	Compatibilidad	Posición	Vocablo	Compatibilidad
1	portátil	0,84886	19	página web	0,762336
2	teclado	0,973801	20	contraseña	0,984674
3	monitor	0,614702	21	e-mail	0,548354
4	torre	0,56671	22	bandeja de entrada	0,444669
5	ratón	0,990916	23	enviar	0,035521
6	cable	0,364104	24	virus	0,030091
7	Wifi	0,243616	25	antivirus	0,002693
8	conexión telefónica	0,009765	26	Emule	0,100015
9	escritorio	0,110538	27	Ares	0,82571
10	papelera de reciclaje	0,011503	28	descarga directa	0,857713
11	Google	0,318346	29	subir	0,067378
12	Tuenti	0,389047	30	doble clic	0,346171
13	Twitter	0,165077	31	Mi PC	0,275321
14	Facebook	0,610439	32	icono	0,28061
15	Messenger	0,107167	33	carpeta	0,096337
16	clic	0,03084	34	ventana	0,190164
17	campus virtual	0,187384	35	Windows Vista	0,174066
18	(Microsoft) Word	0,171603			

Tabla 6.13. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica. Informante 016. Centro de interés 01. *Partes del cuerpo humano*

En inglés, el informante 87 comienza aportando palabras con un alto índice de compatibilidad pero a partir de un determinado momento los procesos de reentrada al núcleo son muy leves. Al igual que ocurre con otros hablantes que se desvían con palabras descentralizadas, vuelve a conectar con el prototipo en varias ocasiones con menor intensidad que en las primeras secuencias léxicas. Por ejemplo, aunque parezca que el informante se ha desviado totalmente del prototipo a partir de la palabra *tool*, no es así porque vuelve a entrar en el núcleo con una palabra que forma parte del nivel 6 del listado de palabras nucleares, (*Microsoft*) *Word*.

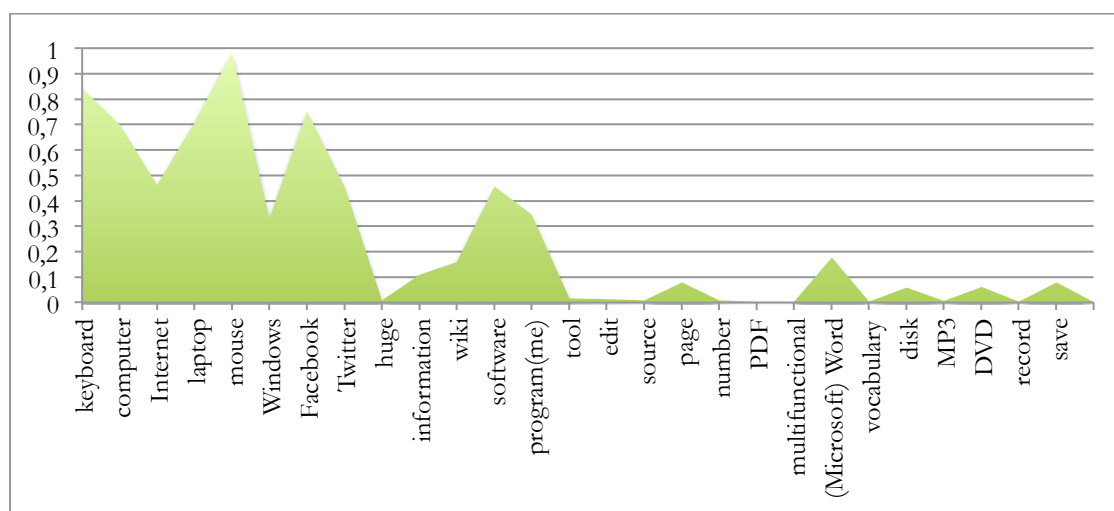


Gráfico 6.32. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado con desviación lineal. Informante 087. Centro de interés 09. *Computer and Internet*

En francés, se ha seleccionado el informante que ha aportado más palabras, concretamente el 30, que destaca por su riqueza léxica y por la organización de las unidades léxicas en varias secuencias. Para comenzar, el informante repite las dos palabras que componen la denominación del centro de interés (*ordinateur* e *internet*), da cuenta de la importancia que tiene para él en ese centro *jeu vidéo* y se refiere a dos partes del ordenador (*câble* y *écran*). A continuación, aporta el vocablo genérico *réseau social*, menciona tres nombres propios en ese ámbito (*Facebook*, *Tuenti* y *Twitter*) y proporciona más información sobre lo que se puede realizar (*forum*, *parler* y *chat*). Finalmente, realiza asociaciones por pares (*problème-technicien*, *Spotify-musique*, *jeu-jeu vidéo*, etc.) con un índice de compatibilidad inferior en la mayoría de los casos.

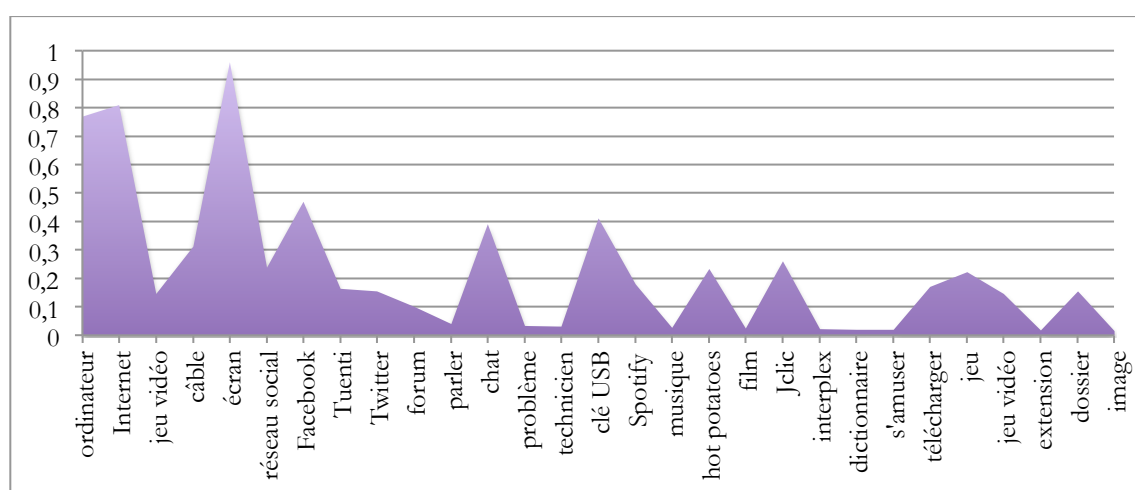


Gráfico 6.33. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado con desviación lineal. Informante 034. Centro de interés 09. *Ordinateurs et Internet*

Cabe destacar, que sobre todo en este centro, es frecuente encontrar en los listados en inglés y francés vocablos que se refieren propiamente a la nomenclatura del núcleo, como *ordenador* e *internet*. De forma inconsciente el informante intenta volver a conectar directamente con el tema, lo que le lleva a repetir la palabra que hace referencia al centro de interés. Quizás este mismo proceso lo realiza el informante en español sin necesidad de plasmarlo en papel porque reacciona de forma más rápida y no lo escribe. También se observa que cuando el informante realiza la prueba en lengua extranjera incorpora más nombres propios que en español, quizás porque carece del vocabulario necesario.

6.5. Análisis de las relaciones semánticas a través del DispoGrafo

Los investigadores chilenos Echeverría (2001, 2002) y Ferreira (2006) hacen referencia al modelo conexionista para explicar la organización de las palabras en la mente. Esta teoría fue desarrollada por Quillian (1968) y parte de la premisa de que una red semántica (también denominada grafo) está formada por conceptos o nodos con unas características específicas que se unen mediante arcos con los cuales comparten información. De esta manera, el significado de una palabra o nodo no puede explicarse de forma aislada sino en relación con la red de la que forma parte. Concretamente Echeverría (2001: 91, 2002) y Ferreira (2006: 27) mencionan como trabajo de partida el modelo conexionista distribuido de Rumelhart y McClelland (1986) que propone que la memoria se basa en redes neuronales entendidas como teorías computacionales complejas cuyo funcionamiento es distribuido y en paralelo.

A través del programa informático DispoGrafo, es posible conocer las redes semánticas que se organizan en la mente de los hablantes cuando realizan la prueba de disponibilidad léxica. Esas redes pueden ser dibujadas mediante la creación de grafos de un determinado centro de interés o subgrafos generados a partir de una palabra concreta. La especificidad de las redes vendrá marcada por una mayor o menor exigencia de la poda o selección realizada en dos niveles: sobre el número de conexiones o aristas y sobre el peso o fuerza de cada una de ellas. Para completar los resultados y analizar más en profundidad la asociación de las palabras, se han exportado listados con indicación al número de conexiones entre los diferentes pares

de palabras siguiendo el orden de aparición.

Los análisis generados tratan de responder a varias hipótesis de investigación iniciales. Por un lado, se pretende corroborar la diferente naturaleza de los centros de interés según las descripciones realizadas por Hernández Muñoz (2006: 150-156) y Sánchez Saus-Laserna (2011: 196-209) y su incidencia en los resultados de cada centro. Por otro lado, se realizan comparaciones léxicas intramuestrales entre los tres idiomas del estudio para desgranar las principales semejanzas y diferencias en la organización del léxico disponible.

6.5.1. Análisis de las relaciones semánticas según el centro de interés

A modo de presentación, en este apartado se destaca la información significativa sobre las redes semánticas en cada centro de interés que ayudarán a explicar los grafos generados mediante el DispoGrafo. En primer lugar, se muestran datos generales como el número de palabras y vocablos y el índice de densidad léxica analizados en el capítulo 4 que complementan los datos sobre la clasificación de los centros de interés. En segundo lugar, se incluye el número de vocablos considerados nucleares, es decir, aquellos resultantes de los seis niveles generados a partir de la herramienta *Fuzzy Expected Value*²⁹³.

Para conocer el grado de palabras nucleares que el informante aporta con respecto al total de vocablos, se ha calculado la densidad de prototipicidad. Este indicador se obtiene al dividir el número total de vocablos por el número de vocablos considerados centrales o nucleares. Cuanto mayor sea el número la densidad de prototipicidad quiere decir que más palabras forman parte del grupo de palabras nucleares con respecto al total.

$$\text{Densidad de prototipicidad}^{294} = \frac{\sum V}{VN}$$

En el caso del centro 1, *Partes del cuerpo humano*, de cada 5,674 palabras que menciona un informante, es decir entre 5 y 6 palabras que escribe, menciona un

²⁹³ Para más información sobre el cálculo de este índice, consúltase el apartado 2.5.2.

²⁹⁴ Donde $\sum V$ es la suma de vocablos y VN los vocablos nucleares.

vocablo nuclear. En cambio, en el noveno centro, *Ordenadores e internet*, la proporción es bastante mayor, siendo uno de cada 8 vocablos.

En último lugar, se incluye el número de aristas extraídas del DispoGrafo con independencia de su peso. Es decir, el resultado contabiliza el número de conexiones diversas sin tener en cuenta el número de veces que se repiten. Se ha calculado la densidad de conectividad dividiendo el número de vocablos por el número de conexiones. Si el número de conexiones es mayor, significa que la densidad será menor, porque las relaciones entre las palabras no están tan definidas y viceversa.

$$\text{Densidad de conectividad}^{295} = \frac{\sum C}{\sum V}$$

En líneas generales, los datos numéricos tratan de corroborar la diferente ontogenia de los centros de interés. Los centros considerados prototípicos, como el caso de *Partes del cuerpo humano*, presentan un léxico más limitado en número de vocablos pero más productivo en número de palabras. Esa característica se ve reflejada en la densidad léxica, en el número de conexiones por vocablos pero, también, en la densidad de prototipicidad y de conectividad con valores superiores en otros centros. Al ser un centro muy productivo con un número reducido de vocablos, las conexiones son más sólidas.

En el lado opuesto se situarían los centros más heterogéneos o relacionales como es el caso del centro 9, referido a *Ordenadores e internet*. Este campo léxico está compuesto por muchas lexías diferentes con un número de palabras inferior a la media. Ello conlleva que ocupe el último rango en el índice de densidad léxica en español e inglés y el penúltimo en francés. Asimismo, presenta tan solo 4,06 conexiones por vocablo en español, 2,92 en inglés y 2,11 en francés por lo que se sitúa en los niveles más bajos de los listados presentes en los tres idiomas (Tabla 6.14, Tabla 6.15 y Tabla 6.16).

CI	Palabras totales	Vocablos	Vocablos nucleares	Densidad de prototipicidad	Conexiones	Densidad de conectividad
1	4493	244	43	5,674	2031	8,32
2	3759	238	41	5,805	1780	7,48
3	4565	530	73	7,260	3059	5,77
4	3718	392	50	7,840	1965	5,01
5	3711	483	67	7,209	2527	5,23

²⁹⁵ Donde $\sum C$ se refiere a la suma de conexiones diferentes y $\sum V$ a la suma de vocablos.

6	2641	190	28	6,786	1179	6,21
7	3040	469	64	7,328	2071	4,41
8	3600	478	70	6,829	2576	5,39
9	3290	555	69	8,043	2256	4,06
Media	3646,33	397,67	56,11	7,087	2160,44	5,76

Tabla 6.14. Resultados comparativos en español en relación con palabras, vocablos y conexiones

CI	Palabras totales	Vocablos	Vocablos nucleares	Densidad de prototipicidad	Conexiones	Densidad de conectividad
1	2571	115	26	4,423	1065	4,36
2	2076	151	28	5,393	1022	4,29
3	2894	287	50	5,740	1817	3,43
4	2120	244	32	7,625	1126	2,87
5	2344	343	54	6,352	1634	3,38
6	1675	165	26	6,346	778	4,09
7	1909	310	45	6,889	1295	2,76
8	1928	280	44	6,364	1321	2,76
9	2251	460	54	8,519	1620	2,92
Media	2196,44	261,67	39,89	6,560	1297,56	3,43

Tabla 6.15. Resultados comparativos en inglés en relación con palabras, vocablos y conexiones

CI	Palabras totales	Vocablos	Vocablos nucleares	Densidad de prototipicidad	Conexiones	Índice de conectividad
1	427	78	22	3,545	314	0,282
2	365	80	22	3,636	272	0,275
3	452	166	37	4,486	388	0,223
4	323	113	25	4,520	258	0,221
5	389	135	32	4,219	342	0,237
6	249	53	16	3,313	161	0,302
7	282	125	25	5,000	239	0,200
8	299	119	28	4,250	261	0,235
9	316	127	28	4,536	268	0,220
Media	344,67	110,67	26,111	4,238	278,111	0,193

Tabla 6.16. Resultados comparativos en francés en relación con palabras, vocablos y conexiones

Por último, se presentan en forma de tabla con los cliques y *clusters* más frecuentes en lengua española. El peso de los *clusters* difiere según la naturaleza del centro de interés, contando con más peso en los centros más compactos y viceversa. Se podría decir que un buen método de enseñanza de español empezaría insistiendo en las palabras que presentan más conexiones y, por tanto, que serán las que generen redes semánticas de mayor solidez. A modo de ejemplo, si los vocablos *pizarra-tiza* aparecen conectados hasta 48 veces supone que las relaciones entre ambos están muy afianzadas. El alumno que aprenda una de las palabras necesitará en algún momento utilizar la otra y, por tanto, la enseñanza de ambas simultáneamente ayuda a generar núcleos semánticos similares a los de un hablante nativo.

CI	N	Cliques	T ¹	N	Cluster 2	T	N	Cluster 3	T	N	Cluster 4	T	N	Cluster 5	T	N	Más de 5 clusters	T
1	45	mano -> dedo	639	63	nariz, boca	584	43	ojo, nariz, boca	497	28	oreja, ojo, nariz, boca	254	8	cabeza, oreja, ojo, nariz, boca	110	5	cabeza, pelo, oreja, ojo, nariz, boca	81
	43	nariz-> boca		59	mano, dedo		32	oreja, nariz, boca		14	pierna, pie, tobillo, rodilla		8	pelo, oreja, ojo, nariz, boca				
	39	dedo -> uña		52	pierna, pie		20	ojo, pestaña, ceja		10	mano, dedo, brazo, codo		7	ojo, nariz, boca, diente, lengua				
2	53	camisa - >camiseta	565	66	camisa, camiseta	501	35	sujetador, calzoncillo, braga	403	11	sujetador, calzoncillo, braga, tanga	189	5	camisa, camiseta, falda, pantalón, vestido	59	4	camisa, camiseta, blusa, falda, pantalón, vestido, (pantalón) vaquero	12
	27	falda -> vestido		51	calzoncillo, braga		17	calzoncillo, braga, tanga		8	sujetador, calzoncillo, braga, calcetín		4	camisa, camiseta, blusa, falda, vestido				
	27	braga -> calzoncillo		45	falda, pantalón		16	falda, pantalón, vestido		8	falda, pantalón, vestido, (pantalón) vaquero		4	jersey, camisa, camiseta, falda, pantalón				
3	37	carne -> pescado	578	47	pescado, carne	611	10	ensalada, lechuga, tomate	297	5	pescado, fruta, carne, verdura	81	3	pescado, fruta, carne, verdura, legumbre	20	2	pescado, arroz, pasta, fruta, carne, verdura	6
	24	lechuga -> tomate		42	lechuga, tomate		10	lechuga, tomate, patata		4	agua, vino, cerveza, leche		2	jamón, pollo, ternera, cerdo , cordero				
	20	cerveza -> vino		37	vino, cerveza		10	pollo, ternera, cerdo		4	agua, vino, cerveza, refresco		2	sandía, melón, piña, melocotón, uva				
4	74	mesa-silla 74	468	98	silla, mesa	451	18	pizarra, tiza, borrador	222	6	regla, escuadra, cartabón, compás	51	2	pizarra, pizarra digital, ordenador, (ordenador) portátil, internet	10	2	silla, mesa, pupitre, pizarra, tiza, borrador	2
	50	pizarra-tiza		82	lápiz, boli(grafo)		15	mesa, silla, pizarra		6	mesa, silla, pizarra, tiza		2	silla, pizarra, tiza, boli(grafo), lápiz				
	48	lápiz-bolígrafo		56	pizarra, tiza		10	boli(grafo), lápiz, goma (de borrar)		4	libro, libreta, lápiz, boli(grafo)		2	silla, mesa, pupitre, pizarra, tiza				
5	20	calle-avenida	477	30	edificio, casa	505	9	plaza, calle, avenida	195	3	semáforo, moto(cicleta), bicicleta, tráfico	34	2	semáforo, coche, carretera, calle, tráfico	7	3	semáforo, coche, moto(cicleta), bicicleta, (auto)bus, tráfico	4
	18	edificio-casa		29	cine, teatro		9	bar, pub, discoteca		3	restaurante, bar, pub, discoteca		2	semáforo, coche, moto(cicleta), bicicleta, edificio				

²⁹⁶ La letra *T* se refiere al total de conexiones y la letra *N* al número concreto de cliques o *clusters* iguales al que hace alusión.

	18	cine-teatro		24	restaurante, bar		8	edificio, casa, calle		2	semáforo, casa, (auto)bus, metro		2	semáforo, coche, moto(cicleta), bicicleta, tráfico				
6	35	bici(cleta), moto(cicleta)	343	66	bici(cleta), moto(cicleta)	326	29	coche, bici(cleta), moto(cicleta)	261	15	coche, bici(cleta), moto(cicleta), (auto)bus	155	6	coche, bici(cleta), avión, moto(cicleta), tren	78	7	coche, bici(cleta), avión, moto(cicleta), tren, (auto)bus	71
	31	coche, moto(cicleta)		46	avión, tren		22	metro, tranvía, tren		10	coche, avión, tren, (auto)bus		6	coche, bici(cleta), avión, moto(cicleta), (auto)bus				
	31	moto(cicleta), bici(cleta)		43	avión, barco		20	avión, tren, (auto)bus		9	coche, bici(cleta), avión, moto(cicleta)		6	bici(cleta), moto(cicleta), metro, tren, (auto)bus				
7	48	fútbol, baloncesto	326	62	fútbol, baloncesto	355	23	fútbol, tenis, baloncesto	151	7	parchís, ajedrez, oca, damas	53	3	fútbol, tenis, balonmano, baloncesto, voleibol	15	2	parchís, (juego de) cartas, oca, cine, teatro, concierto	7
	26	parchís, oca		35	parchís, oca		16	fútbol, balonmano, baloncesto		6	fútbol, tenis, balonmano, baloncesto		2	parchís, (juego de) cartas, oca, veo-veo, corro de la patata				
	17	cine, teatro		27	cine, teatro		10	parchís, (juego de) cartas, oca		5	ajedrez, fútbol, tenis, baloncesto		2	parchís, ajedrez, oca, damas, dominó				
8	27	profesor, maestro	415	39	profesor, maestro	477	6	profesor, médico, abogado	170	3	fontanero, carpintero, electricista, albañil	27	2	camarero, azafato, recepcionista, guía turístico, agente	3	2	jardinero, abogado, músico, arquitecto, ingeniero, escritor, enterrador	2
	25	médico, enfermero		33	enfermero, médico		6	carnicero, pescadero, frutero		3	enfermero, médico, policía, bombero		2	enfermero, juez, médico, abogado, notario				
	17	policía, bombero		32	policía, bombero		5	panadero, carnicero, pescadero		2	profesor, enfermero, doctor, maestro		2	albañil, perito, peón, topógrafo, encofrador				
9	43	ratón, teclado	337	75	teclado, ratón	368	45	pantalla, teclado, ratón	165	9	pantalla, teclado, ratón, torre	59	2	pantalla, teclado, ratón, USB, pendrive	14	2	software, hardware, disco duro, cable, pendrive, puerto USB	2
	32	teclado, ratón		51	pantalla, teclado		12	teclado, ratón, portátil		8	pantalla, teclado, ratón, portátil		2	pantalla, teclado, ratón, tecla, cable				
	31	pantalla, teclado		40	pantalla, ratón		11	teclado, ratón, monitor		4	pantalla, teclado, ratón, tecla		2	pantalla, teclado, ratón, portátil, disco duro				

Tabla 6.17. Cliques y *clusters* más frecuentes en lenguas española

6.5.2. Organización del léxico según el idioma

Los análisis realizados muestran de forma gráfica las relaciones semánticas en cada centro de interés y los procesos cognitivos de activación del léxico. A través del estudio de los subcentros y *clusters* detectados es posible conocer qué tipo de fenómeno asociativo ha llevado a la mente de un hablante una palabra y no otra. Ferreira (2006) distingue tres tipos de *priming* o memoria implícita que explican las relaciones léxicas en las encuestas de disponibilidad léxica: *priming semántico*, *priming fonológico* y *priming perceptual*. El *priming* semántico puede ser a corto plazo, cuando una palabra es evocada inmediatamente por la lexía anterior, y a largo plazo cuando los vocablos se encuentran relacionados de forma implícita y se activan en un determinado contexto. El ejemplo proporcionado por el investigador chileno relaciona el *cluster eye-nose-mouth* con un *priming* semántico a corto plazo y el *cluster nine-eleventh, twin-towers* en el centro de interés relacionado con el terrorismo con el *priming* a largo plazo ya que se refiere a un hecho acaecido hace una década pero con un gran impacto.

El *priming* fonológico se produce cuando una palabra activa otra fonológicamente similar, como es el caso de *football-basketball* en inglés. El *priming* perceptual surge cuando la activación de un vocablo se produce por la lectura o producción previa de esa misma palabra. Si la prueba de disponibilidad léxica se realizase justo después de que el hablante haya tomado un café con aspartamo seguramente este segundo vocablo nada frecuente en las listas lo sea para ese individuo en particular. Además, incluye en el análisis el *priming* interlingua que se produce cuando a través de la estructura interna en una lengua se activan palabras en otra lengua. Este fenómeno suele ocurrir cuando la L1 se hace visible en el léxico producido en idioma extranjero lo que muestra que el conocimiento de este último es solo parcial. Por ejemplo, si tras la palabra *glass* el estudiante escribe *lentil* seguramente habrá querido decir en español *gafa-lentilla*, en lugar de *gafa-lenteja*. El segundo vocablo se ha activado en esta ocasión por efecto del *priming* interlingua.

A continuación, se presentan los resultados de tres centros de interés con características diversas: *Partes del cuerpo humano*, *La escuela: muebles y materiales* y

Ordenadores e internet. El primero corresponde al centro más prototípico u homogéneo por antonomasia. El segundo se sitúa en el intermedio de la escala de prototipicidad de los centros y se considera clave para impartir clases en lengua ya que se refiere al universo más cercano que comparten el profesor y el alumno. El tercero corresponde al centro con una estructura más relacional y que no ha sido definido en otros estudios previos.

6.5.2.1. Resultados del centro denominado *Partes del cuerpo humano*

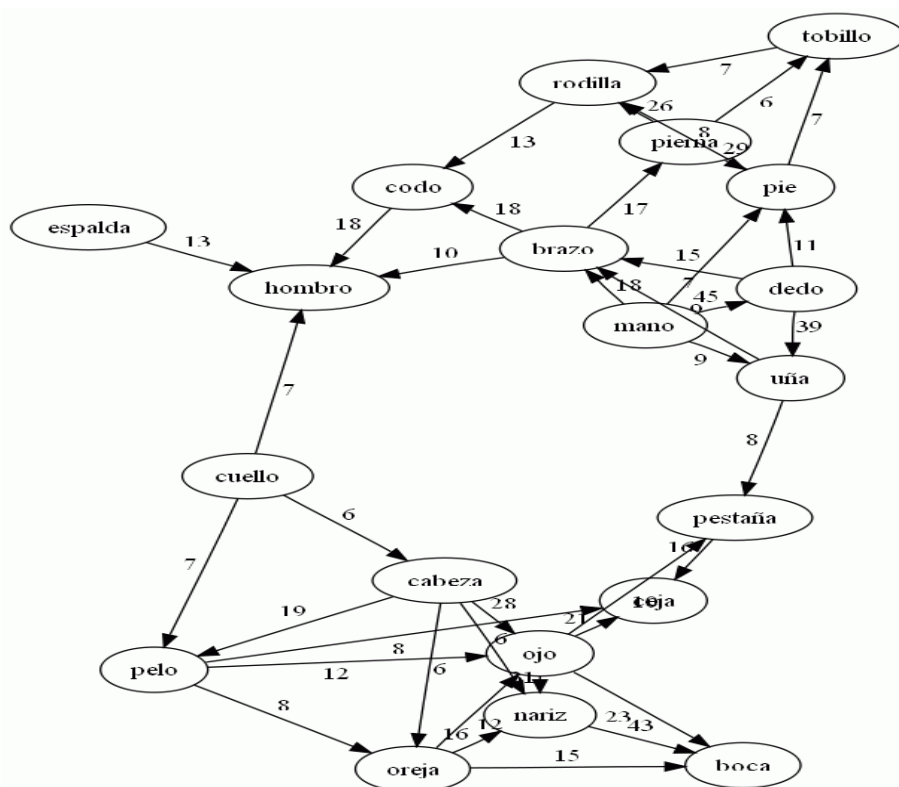
Como se ha abordado en el apartado anterior, el centro de interés referido al cuerpo humano es intrínsecamente compacto puesto que ocupa uno de los primeros rangos según el índice de densidad léxica en la mayoría de los estudios. La relación de la actividad del vocabulario a partir del propio nombre del centro se podría decir que guarda una relación de hiperonimia donde las palabras proporcionadas por el informante aportan una información más específica y subordinada en cierto modo. Los hipónimos suelen estar relacionados directamente con el centro de interés y, por tanto, el número de lexías producidas es menor que en otros ámbitos.

Las conexiones más frecuentes de palabras presentan aristas con un peso de hasta 45 repeticiones, como muestra la relación de meronimia entre *mano* y *dedo*. En la Tabla 6.18 se han incluido los veinte primeros pares de palabras ordenados según el peso de sus aristas. Además de las múltiples relaciones de meronimia presentes en este campo (*dedo-uña*, *pierna-pie*, *boca-diente*), también se observan relaciones que dejan patente un *priming* perceptual tales como *nariz-boca* y *ojo-nariz*.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
mano	->	dedo	45	pie	->	pierna	23
nariz	->	boca	43	nariz	->	oreja	22
dedo	->	uña	39	ojo	->	ceja	21
ojo	->	nariz	31	boca	->	nariz	20
ceja	->	pestaña	31	cabeza	->	pelo	19
pierna	->	pie	29	corazón	->	pulmón	19
boca	->	diente	29	brazo	->	antebrazo	18
cabeza	->	ojo	28	codo	->	hombro	18
rodilla	->	tobillo	26	pie	->	dedo	18
pierna	->	rodilla	26	brazo	->	mano	18
diente	->	lengua	24	brazo	->	codo	18
ojo	->	boca	23	mano	->	brazo	18

Tabla 6.18. Cliques más frecuentes en español. Centro de interés 01. *Partes del cuerpo humano*

En el Grafo 6.1 se pueden observar dos grandes grupos semánticos²⁹⁷. El primero se refiere a partes del cuerpo humano del tronco superior e inferior y el segundo a la cabeza. Los nexos de unión entre ambos grupos son el vocablo *cuello*, que sería la unión física entre el cuerpo y la cabeza, y *pestaña*. En este segundo caso, *pestaña* y *uña* forman parte del *cluster ojo-ceja-pestaña-uña*. Se podría interpretar como partes del cuerpo relacionadas con tratamientos estéticos, de ahí su fuerte relación.



Grafo 6.1. Centro de interés 01. *Partes del cuerpo humano*

En esta misma línea, López González (2014: 56-57) distingue entre dos subcentros o núcleos relacionales: cabeza y extremidades. El primero incluye los merónimos de cabeza (*ojo*, *boca* y *nariz*, entre otros) que incorporan a su vez a sus holónimos (por ejemplo, *boca-labio* y *ojo-pestaña*). El segundo conjunto léxico se organiza en torno a los vocablos relacionados con los órganos de los sentidos. La mayoría de las relaciones se producen por meronimia *mano-dedo-uña* y por proximidad visual de los referentes, *espalda-brazo*.

En lengua extranjera las relaciones producidas siguen procesos semánticos similares de meronimia y de organización espacial. En inglés, la arista con mayor peso la componen *nose-mouth* (39) que ocupa el segundo puesto en lengua española. En este caso, se puede

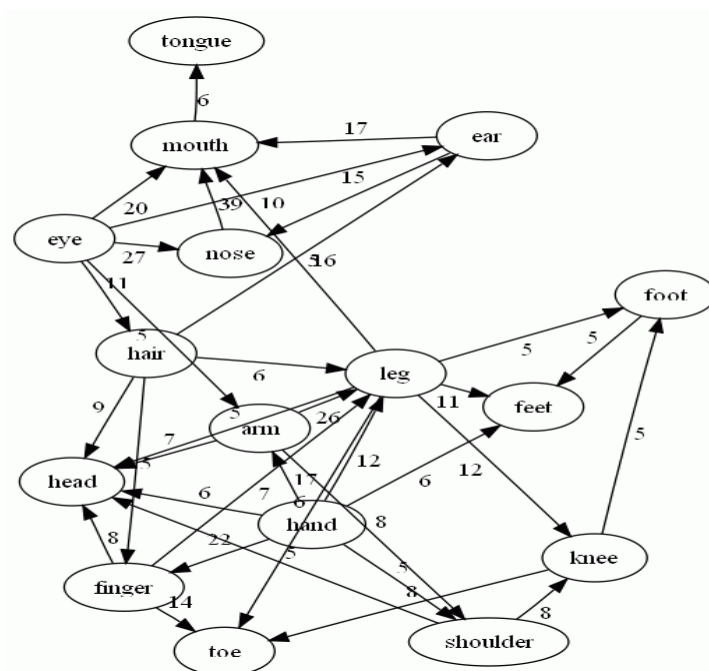
²⁹⁷ En el grafo se han podado las aristas con un peso inferior o igual a cinco y los vocablos con dos o menos relaciones.

haber activado el *priming* fonológico²⁹⁸ de ahí que sitúe también el par *mouth-nose* en sexto lugar.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
nose	->	mouth	39	ear	->	mouth	17
eye	->	nose	27	hand	->	arm	17
leg	->	arm	26	hair	->	ear	16
arm	->	leg	26	head	->	hair	16
hand	->	finger	22	ear	->	nose	15
mouth	->	nose	21	finger	->	hand	15
eye	->	mouth	20	mouth	->	lip	14
finger	->	nail	19	finger	->	toe	14
arm	->	hand	19	arm	->	finger	14
head	->	eye	18	nose	->	ear	14

Tabla 6.19. Tabla 6.20. Cliques más frecuentes en inglés. Centro de interés 01. *Human body parts*

En el Grafo 6.2 los núcleos semánticos no se encuentran tan definidos²⁹⁹. En el subgrafo de la parte superior se encuentran partes del cuerpo relacionadas con la cabeza, unidas a su vez por el vocablo *hair*. En cambio, también está presente en el subgrafo inferior la palabra *head*, por lo que las relaciones no son tan claras: *eye-arm-shoulder-knee*.



Grafo 6.2. Centro de interés 01. *Human body parts*

²⁹⁸ /maʊθ/ y /nəʊz/.

²⁹⁹ En el grafo se han eliminado las aristas con un peso igual o inferior a cuatro y los nodos con dos o menos relaciones.

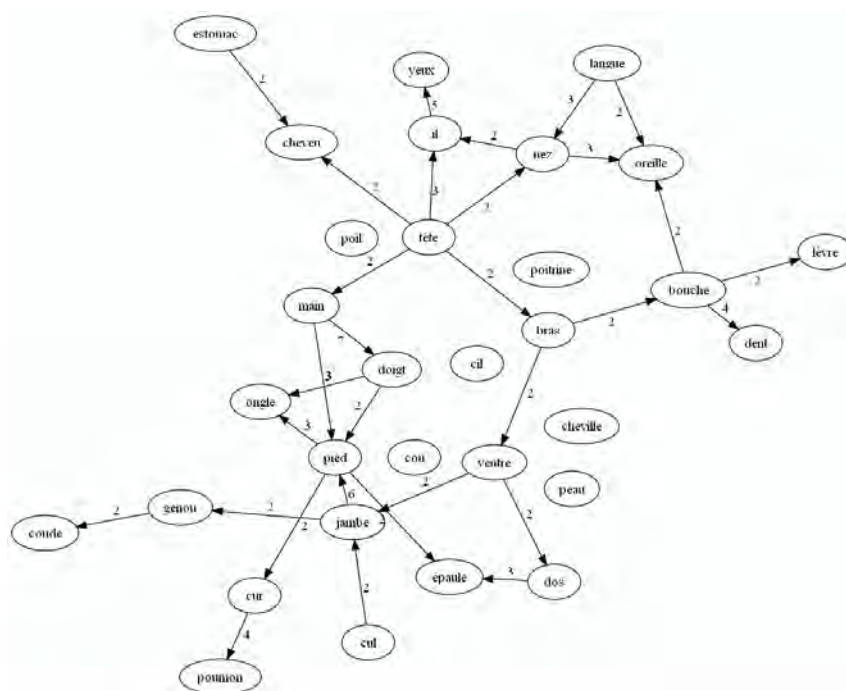
En francés, el número de palabras es inferior al del resto de listas por lo que las relaciones semánticas son menos densas. Cabe mencionar que en francés el plural irregular de *ojo*, *yeux* ha sido precedido hasta en cinco ocasiones de su singular *œil* ya que actúan como vocablos diferentes aunque designen una misma realidad y que se suelen aprender al mismo tiempo.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
ear	->	mouth	17	ear	->	mouth	17
main	->	doigt	7	oreille	->	bouche	3
jambe	->	pied	6	cil	->	sourcil	3
pied	->	jambe	5	tête	->	oeil	3
oeil	->	yeux	5	nez	->	oreille	3
dent	->	langue	4	doigt	->	ongle	3
coeur	->	poumon	4	oreille	->	oeil	3
bouche	->	dent	4	main	->	pied	3
main	->	tête	4	nez	->	bouche	3
pied	->	ongle	3	dos	->	épaule	3

Tabla 6.21. Cliques más frecuentes en francés. Centro de interés 01. *Parties du corps humain*

Los núcleos temáticos y las vínculos no se encuentran tan bien determinados en el Grafo 6.3³⁰⁰ como en español debido a dos factores: el bajo número de informantes y la selección de una lengua extranjera. Esta última característica provoca que ciertos vocablos muestren una relación menos organizada que en español. Así, *estomac* se relaciona en dos ocasiones con *cheveu*; *langue* está unida a *oreille*; y otras palabras más específicas no encuentran relación tales como, *poil*, *poitrine*, *cil* y *peau*, entre otras.

³⁰⁰ En este grafo se han eliminado las aristas con un peso menor o igual a uno y con menos de una relación por vocablo.

Grafo 6.3. Centro de interés 01. *Parties du corps humain*

6.5.2.2. Resultados del centro denominado *La escuela: muebles y materiales*

Uno de los centros clave en la formación lingüística en lengua extranjera del futuro profesorado, es el centro referido a la escuela. Además de los vocablos aportados referidos al mobiliario y al material escolar, existen otros vocablos con los que algunos informantes relacionan la escuela como *amigo*, *aprobar*, *castigo*, *jugar*, *nota*, *parte*³⁰¹, *suspender*, etc. Se podría decir que las unidades léxicas aportadas por los informantes ayudan a definir la percepción de la escuela por parte del futuro profesorado. En este sentido, Fasce, Echeverría, Matus, Ortiz, Palacios y Soto (2009) recogen los vocablos y relaciones semánticas que definen los estudiantes de medicina y médicos en ejercicio como cualidades de la profesión: *excelencia*, *empatía*, *responsabilidad*, entre otras.

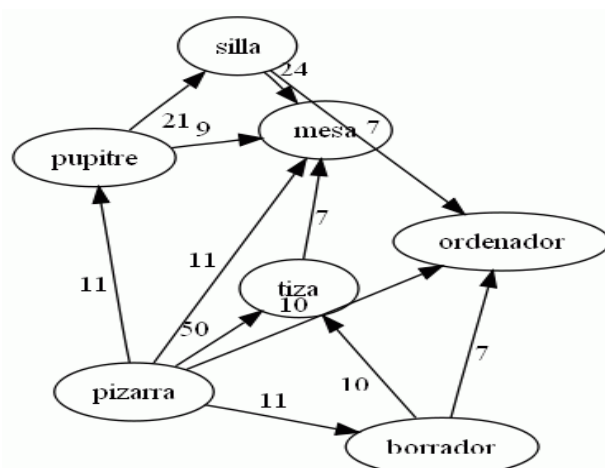
Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
mesa	->	silla	74	pupitre	->	silla	21
pizarra	->	tiza	50	boli(grafo)	->	goma (de borrar)	19
lápiz	->	boli(grafo)	48	silla	->	pizarra	19

³⁰¹ Referido a los informes sobre una conducta inadecuada.

boli(grafo)	->	lápiz	34	goma (de borrar)	->	sacapuntas	18
libro	->	cuaderno	26	mesa	->	pupitre	16
tiza	->	borrador	25	profesor	->	alumno	16
lápiz	->	goma (de borrar)	25	boli(grafo)	->	rotulador	15
silla	->	mesa	24	ventana	->	puerta	15
libreta	->	libro	24	pupitre	->	pizarra	14
libro	->	libreta	24	regla	->	compás	14
escuadra	->	cartabón	24				

Tabla 6.22. Cliques más frecuentes. Centro de interés 04. *La escuela: muebles y materiales*

El Grafo 6.4³⁰² presenta en dos subgrupos la realidad más directa del alumno: *silla* o *pupitre* relacionado con *mesa* y, por otro lado, *tiza*, *pizarra* y *borrador*. Se observa, asimismo, que el vocablo *ordenador* está relacionado tanto con *silla* como con *pizarra* y *borrador* quizás como un recurso tradicional en contraposición con las nuevas tecnologías utilizadas en el mundo de la enseñanza.

Grafo 6.4. Centro de interés 04. *La escuela: muebles y materiales*

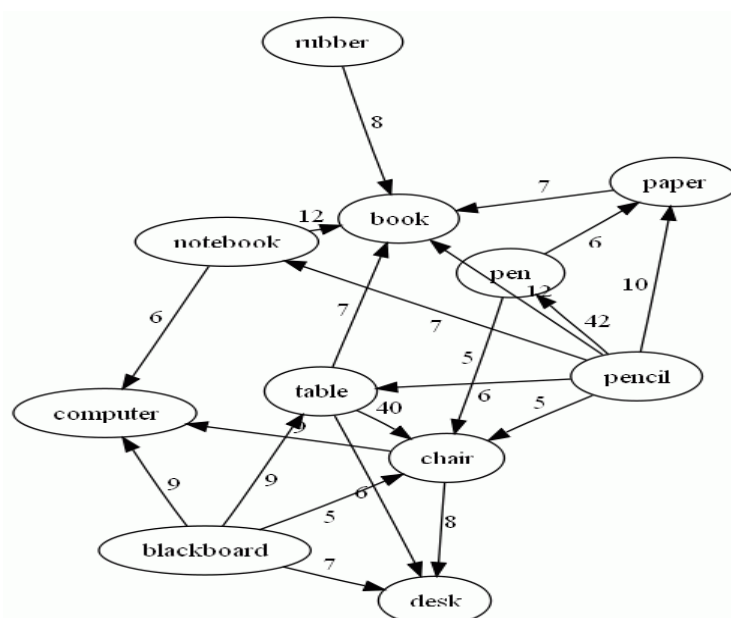
En inglés, las relaciones entre pares de palabras son similares al español. No obstante, algunos *clusters* son más importantes debido al *priming* fonológico, tales como *rubber-ruler* y *pencil-pen*. El Grafo 6.5³⁰³ presenta relaciones más difusas donde se podría interpretar que existe un subgrupo en la parte superior referido a los materiales de escritura y un subgrupo en la parte inferior referido al mobiliario.

³⁰² En este grafo se han eliminado las aristas con un peso menor o igual a cinco y con menos de dos relaciones por vocablo.

³⁰³ En este grafo se han eliminado las aristas con un peso menor o igual a cuatro y con menos de dos relaciones por vocablo.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
pencil	->	pen	42	book	->	workbook	12
pen	->	pencil	41	pencil	->	book	12
table	->	chair	40	table	->	blackboard	11
book	->	notebook	29	book	->	pencil	10
chair	->	table	29	pencil	->	paper	10
chair	->	blackboard	17	rubber	->	ruler	9
desk	->	chair	17	blackboard	->	table	9
door	->	window	13	pen	->	rubber	9
window	->	door	13	chair	->	computer	9
notebook	->	book	12	teacher	->	student	9
blackboard	->	chalk	12	blackboard	->	computer	9

Tabla 6.23. Cliques más frecuentes inglés. Centro de interés 04. *The school: furniture and school material*



Grafo 6.5. Centro de interés 04. *The school: furniture and school material*

En francés, las relaciones más frecuentes se producen entre *tableau-chaise-table*, es decir entre tres vocablos presentes de igual manera en español y en inglés. Como ocurre en el centro referido al cuerpo, los contenidos de este campo léxico suelen aparecer en la mayoría de los manuales de lengua extranjera tanto en Educación Primaria como en Educación Secundaria. No obstante, las asociaciones son menos fuertes que en el primer centro de interés lo que muestra la influencia del estímulo en los resultados de las pruebas de disponibilidad léxica.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
table	->	chaise	10	règle	->	trousse	2
livre	->	cahier	8	stylo	->	ciseaux	2
crayón	->	stylo	5	projecteur	->	ordinateur	2
chaise	->	table	4	tableau	->	gomme	2
stylo	->	crayón	3	gomme	->	règle	2
cahier	->	livre	3	cahier	->	dossier	2
tableau	->	chaise	3	chaise	->	fauteuil	2
table	->	tableau	3	fenêtre	->	porte	2
chaise	->	fenêtre	3	cahier	->	stylo	2
tableau	->	table	2	élève	->	professeur	2

Tabla 6.24. Cliques más frecuentes en francés. Centro de interés 04. *L'école: meubles et matériel*

El Grafo 6.6 generado en esta ocasión solo ha permitido la eliminación de las aristas con un peso menor o igual a dos y con al menos una relación por vocablo. Esa misma criba se realizó en el primer centro pero contaba con más relaciones entre sí, lo que refleja que en este centro el vocabulario está menos cohesionado. Sin embargo, es interesante considerar este grafo como un *cluster* que el alumnado retiene y que además es de gran utilidad en lengua extranjera ya que designa a tres objetos de la clase imprescindibles: la pizarra, la silla y la mesa.

Grafo 6.6. Centro de interés 04. *L'école: meubles et matériel*

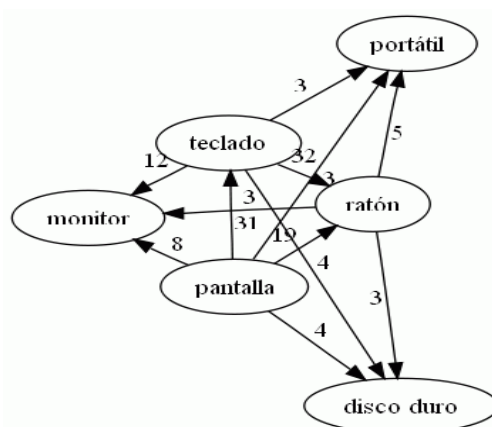
6.5.2.3. Resultados del centro denominado *Ordenadores e internet*

En este campo léxico, el podado de los grafos es considerablemente menor ya que las palabras se encuentran menos conectadas y, por tanto, al descartar las aristas con una frecuencia igual o menor a cinco, cuatro o tres se eliminan todas las palabras del centro de interés. Los vocablos nucleares presentan una gran cohesión en el grafo mientras que el resto se organiza de forma desigual. Los tres primeros pares de palabras están formados por las tres lexías que pertenecen al nivel 1 según su nuclearidad: *ratón, teclado y pantalla*.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
ratón	->	teclado	43	pantalla	->	torre	12
teclado	->	ratón	32	teclado	->	monitor	12
pantalla	->	teclado	31	red social	->	Facebook	11
hardware	->	software	22	ratón	->	impresora	10
ratón	->	pantalla	21	virus	->	antivirus	9
teclado	->	pantalla	20	correo (electrónico)	->	red social	8
pantalla	->	ratón	19	pantalla	->	monitor	8
impresora	->	escáner	17	portátil	->	teclado	8
Facebook	->	Tuenti	16	escáner	->	impresora	8
software	->	hardware	15	página web	->	correo (electrónico)	8
CD(-ROM)	->	DVD	13	archivo	->	carpeta	8
Tuenti	->	Facebook	12				

Tabla 6.25. Cliques más frecuentes en español. Centro de interés 09. *Ordenadores e internet*

Se podría decir que el Grafo 6.7³⁰⁴ crea un núcleo central compuesto por *teclado*, *pantalla*, *ratón* y *monitor* conectado con la palabra *portátil* y *disco duro*. Además de las unidades léxicas relativas a los ordenadores, son reseñables las conexiones entre nombres de redes sociales como *Facebook* y *Tuenti*³⁰⁵ y entre el genérico *red social* y el concreto *Facebook* (once conexiones). Entre los *clusters* más frecuentes cabe mencionar los préstamos *software-hardware* y *CD(-ROM)-DVD*.

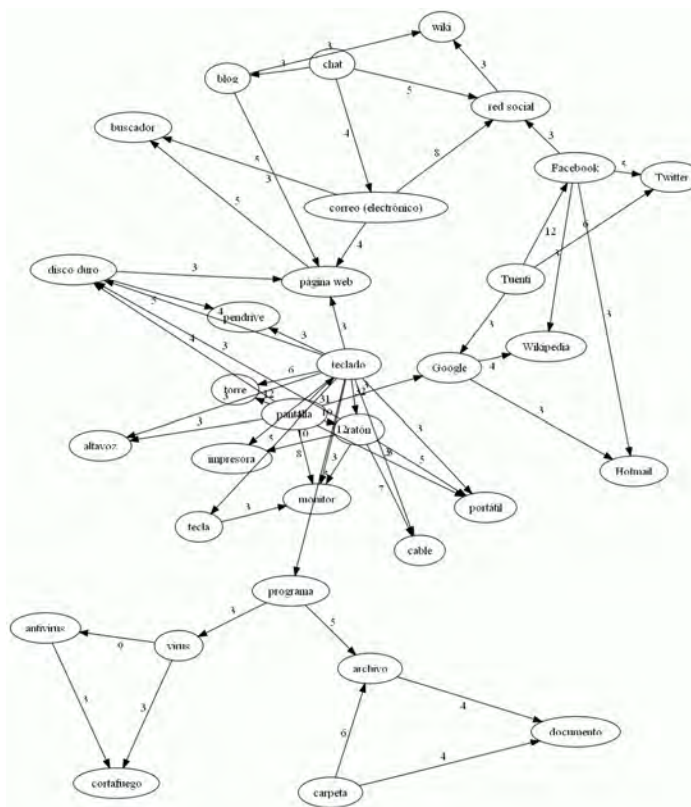


Grafo 6.7. Centro de interés 09. *Ordenadores e internet*

³⁰⁴ En este grafo se han eliminado las aristas con un peso menor o igual a tres y con al menos dos relaciones por vocablo.

³⁰⁵ De *Facebook* a *Tuenti* la arista tiene un peso de 16 y de *Tuenti* a *Facebook* se establecen 12 conexiones.

En la parte inferior, se destacan dos subgrafos generados a partir de la palabra *programa*. En la parte izquierda, los vocablos hacen referencia a la protección de un ordenador (*virus*, *antivirus*, *cortafuegos*) y en la parte derecha a nombres relacionados con operaciones básicas (*archivo*, *documento* y *carpeta*). El peso de las aristas también se ve reflejado en la distribución de los nodos que se encuentran más unidos en el centro donde hay aristas que superan un peso de 30 y más alejados en esta parte inferior y superior donde se recogen aristas con un peso considerablemente menor.



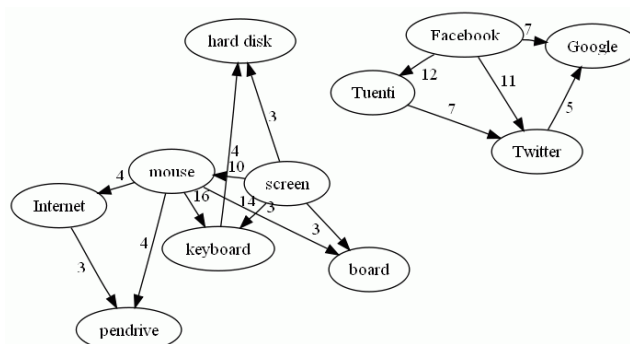
Grafo 6.8. Centro de interés 09. *Ordenadores e internet* (ejemplo ampliado)

En inglés, aparecen como conexiones más frecuentes *software-hardware* en ambas direcciones. En lengua española son préstamos que forman parte del léxico nuclear lo que explica que también en inglés sean palabras nucleares. Sin embargo, el *cluster download-upload* se sitúa entre los diez primeros en inglés pero no en español. Esto puede deberse a que esas palabras son a menudo mantenidas en inglés en los diferentes programas y páginas webs pero en español se utiliza la traducción *descargar* y *subir*. Por otra parte, los nombres propios de redes sociales y buscadores, *Facebook*, *Twitter*, *Tuenti* y *Google* tienen mayor relevancia debido probablemente a que el informante suple la falta de vocabulario en este ámbito con esos nombres propios.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2
software	->	hardware	22	Facebook	->	Twitter
hardware	->	software	22	screen	->	mouse
mouse	->	screen	20	Tuenti	->	Facebook
mouse	->	keyboard	16	Tuenti	->	Twitter
screen	->	keyboard	14	CD(_ROM)	->	DVD
keyboard	->	mouse	14	Facebook	->	Google
download	->	upload	14	hard disk	->	floppy disk
keyboard	->	screen	13	Twitter	->	Google
Facebook	->	Tuenti	12	program(me)	->	virus
printer	->	scanner	12	Twitter	->	Facebook

Tabla 6.26. Cliques más frecuentes en inglés. Centro de interés 09. *Computer and Internet*

El Grafo 6.9 ha sido generado con los vocablos que tienen al menos dos aristas y cuyas conexiones tienen un peso mínimo de tres conexiones, por lo que la delimitación es menor que en los centros anteriores. Se pueden observar dos subcampos independientes: el primero referido a vocablos genéricos (*hard disk*, *screen*, *keyboard*, *board*, *mouse*, *Internet* y *pendrive*) y el segundo a nombres de redes sociales (*Facebook*, *Tuenti* y *Twitter*) y del principal buscador (*Google*).



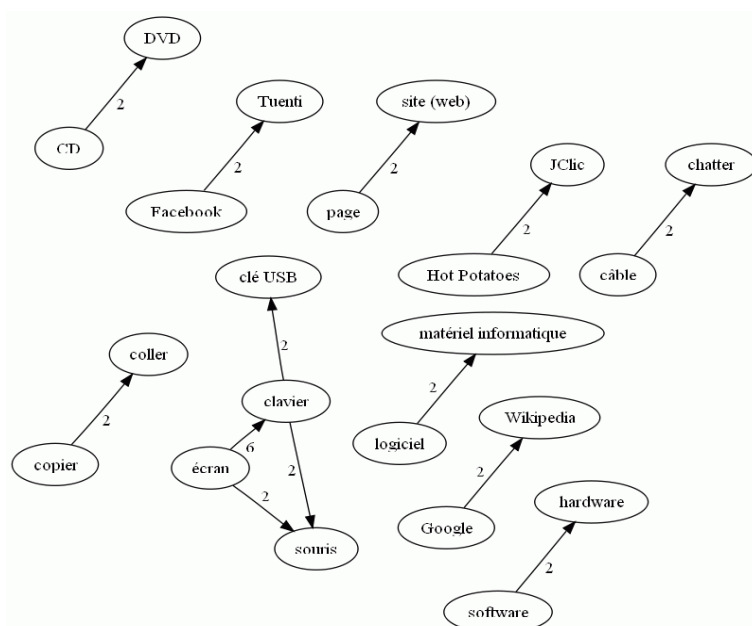
Grafo 6.9. Centro de interés 09. *Computer and Internet*

En francés, debido la influencia anglosajona en el campo de las nuevas tecnologías, existen *clusters* provenientes de préstamos ingleses, tales como los vocablos *software*, *hardware*, *CD*, *DVD* y *page web*. Asimismo, se incluyen nombres propios como *Google*, *Wikipedia*, *Facebook* y *Tuenti*, *Hot Potatoes* y *JClic*. Estos dos últimos corresponden a aplicaciones informáticas diseñadas para la creación de material educativo por lo que la activación de esos vocablos podría dar buena cuenta de que el futuro profesorado conoce recursos informáticos para aplicarlo en la enseñanza de su materia.

Nodo 1	Dirección	Nodo 2	Peso	Nodo 1	Dirección	Nodo 2	Peso
écran	->	clavier	6	Facebook	->	Tuenti	2
souris	->	écran	5	copier	->	coller	2
souris	->	clavier	4	clavier	->	clé USB	2
Hot Potatoes	->	JClic	2	CD	->	DVD	2
écran	->	souris	2	logiciel	->	matériel informatique	2
software	->	hardware	2	Google	->	Wikipedia	2
page	->	site (web)	2	disquette	->	page web	2
page web	->	souris	2	clavier	->	CD	1
clavier	->	souris	2	CD	->	souris	1
câble	->	chatter	2	souris	->	réseau social	1

Tabla 6.27. Cliques más frecuentes en francés. Centro de interés 09. *Ordinateurs et Internet*

El Grafo 6.10 contempla todos aquellos vocablos que tienen al menos una arista y cuyo valor es de un mínimo de dos conexiones. A simple vista se observa que la mayoría de los nodos corresponden a *clusters* entre pares salvo la relación *souris-écran-clavier-clé USB*. Los tres primeros términos corresponden a lexías con un índice elevado nuclearidad y presentes en los tres idiomas. Los *clusters* son muy similares en inglés y en español dada la universalidad del lenguaje informático y de los recursos y páginas webs: *hardware-software*, *Google-Wikipedia*, *CD-DVD*, entre otras.

Grafo 6.10. Centro de interés 09. *Ordinateurs et internet*

Tras analizar las redes semánticas en los tres idiomas, se advierte que los campos y subcampos están mejor delimitados y marcados en la lengua materna de los informantes, es decir, en español. En inglés y francés el número de palabras aportadas por los informantes es inferior por lo que a su vez el número de relaciones también es menor. Para compensar ese desequilibrio, el podado realizado en cada idioma ha sido diferente de forma que se han mostrado las conexiones más frecuentes en cada lengua con indicación al grado de exigencia incluido en el análisis realizado en el DispoGrafo.

6.6. Hacia una propuesta didáctica

La disponibilidad léxica está asociada desde sus orígenes a la selección del léxico para la enseñanza de la lengua cuando los investigadores Gougenheim *et al.*, (1964) elaboran un vocabulario elemental para la enseñanza del francés a extranjeros, inmigrantes y a los habitantes de la Unión Francesa. La mayoría de estudios posteriores destacan la importancia del léxico disponible en la enseñanza y adquisición de una lengua materna o extranjera pero se limitan a anexar el diccionario de léxico disponible sin aportar ninguna propuesta concreta. No obstante, existen trabajos que incluyen aportaciones didácticas y metodológicas e incluso un programa informático.

Como fruto de un proyecto de investigación *Metacognición en lengua materna. Software multimedial para su desarrollo* (1997-1999), Alba y Echeverría elaboran varios programas informáticos para la enseñanza del español. Uno de ellos se titula *Vocabulario Disponible* y tiene como objetivo la enseñanza del léxico disponible recogido en un estudio realizado en el año 1999 a más de 2000 estudiantes chilenos. Además de aportar una fundamentación teórica, el software permite una evaluación previa de la capacidad léxica del estudiante para que conozca la situación de partida antes de empezar a realizar las actividades³⁰⁶.

Otro de los trabajos que llevan a la práctica docente los resultados del léxico disponible es la tesis doctoral de García Casero (2013). La muestra está formada por estudiantes de 4.º de la ESO, etapa educativa donde la autora es docente. Tras realizar análisis comparativos en el plano cualitativo y cuantitativo, presenta una propuesta didáctica destinada a mejorar la competencia léxica de los estudiantes de la Educación Secundaria Obligatoria. Las 16 actividades propuestas vienen acompañadas de unas orientaciones metodológicas y una fundamentación teórica que le sirven de guía al profesorado tanto de español como lengua materna como de español como lengua extranjera.

El objetivo de este apartado es describir la aplicación práctica de los conceptos teóricos y análisis previos sobre el léxico nuclear y las relaciones de las palabras. En primer lugar, se presentan los criterios más usuales de selección léxica en los estudios de disponibilidad. En segundo lugar, se destaca la importancia de los aspectos socioculturales en el léxico disponible para reflexionar sobre la selección del vocabulario en función del objetivo de aprendizaje: español como segunda lengua en una ciudad española, español para extranjeros, español para un futuro profesorado, etc.

En tercer lugar, se comparan los centros de interés tradicionales con las nociones específicas del Plan Curricular del Instituto Cervantes (PCIC) para comprobar el alcance temático de los centros. A continuación, se evalúa posible presencia de matices culturales a través del análisis de los vocablos de un mismo

³⁰⁶ Para más detalles, consulte el apartado 2.6 o el artículo de Echeverría (2002) disponible en http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342002005100012&lng=en&nrm=iso&ignore=.html [Fecha de última consulta el 5 de marzo de 2015].

centro en los tres idiomas del estudio. Para completar la valoración del léxico propuesto, se analiza el listado del léxico nuclear de un centro presente en el inventario de nociones específicas, *Comidas y bebidas*, con los niveles del Marco Común Europeo de Referencia para las Lenguas.

En cuarto y último lugar, se incluye una propuesta práctica que consiste en la creación de una base de datos disponible en línea para los profesionales de la enseñanza de idiomas. En ese apartado se destacan los aspectos teóricos como la importancia de conocer las relaciones más frecuentes entre las palabras (con ejemplos extraídos mediante el programa DispoGrafo) y la clasificación de los vocablos en niveles según el índice de compatibilidad.

6.6.1. Selección léxica según diferentes criterios de corte

Desde el nacimiento de la disponibilidad léxica en los años cincuenta en Francia, se hace patente la necesidad de elaborar unas listas de vocabulario para enseñar a la población francés. Por ello, numerosos estudios establecen diversos índices para obtener el léxico más disponible (Mackey, 1971a; Samper Padilla, 1999; Carcedo González, 2000a; Ávila Muñoz y Villena Ponsoda, 2006, 2010; etc.).

Uno de los investigadores que se cuestionan la idoneidad de incluir todo el vocabulario de las encuestas es Mackey (1971a: 39). A partir de un estudio realizado en la región canadiense de Acadia, el autor analiza cuáles son los vocablos escritos por un solo informante. En este caso, de los 10 000 vocablos presentes en los dieciséis centros de interés tradicionales, más del 44 % habían sido aportados por una persona.

En el estudio actual los datos son similares puesto que en los tres idiomas más del 42 % de las palabras han sido actualizadas por un solo informante. La diferencia entre un idioma a otro puede deberse a la influencia del número de informantes en cada caso. Se observa que a mayor número de informantes menor proporción de vocablos aportados una sola vez. Los vocablos aportados por un único sujeto son fácilmente identificables gracias al porcentaje de aparición. Esas palabras en español cuentan con un porcentaje de 0,58 %, en inglés 0,67 % y en francés 4,76 %. No

obstante, no procede una comparación de los tres idiomas debido a que el número de informantes en español es de 171, en inglés 150 y en francés 21.

CI	Español			Inglés			Francés		
	Total	UI	%	Total	UI	%	Total	UI	%
C1	244	69	28,28	115	38	33,04	78	29	37,18
C2	238	84	35,29	151	58	38,41	80	32	40
C3	530	207	39,06	287	117	40,77	166	84	50,6
C4	392	192	48,98	244	112	45,9	113	64	56,64
C5	483	220	45,55	343	169	49,27	135	61	45,19
C6	190	62	32,63	165	85	51,52	53	26	49,06
C7	469	207	44,14	310	134	43,23	125	75	60
C8	478	192	40,17	280	128	45,71	119	66	55,46
C9	555	286	51,53	460	253	55	127	72	56,69
Total	3579	1519		2355	1094	402,85	996	509	450,82
Media	397,67	168,78	42,44	261,67	121,56	44,76		56,5	50,09
								6	

Tabla 6.28. Vocablos aportados por un único informante y su proporción en el centro de interés con respecto al total de vocablos en lengua española

A continuación, se analiza la selección del léxico siguiendo criterios más complejos. Por un lado, se tiene en cuenta el índice de disponibilidad, la frecuencia de acumulada y el porcentaje de aparición de cada vocablo. Por otro lado, se establece el criterio de determinación del léxico relevante a partir de la teoría de los conjuntos difusos aplicada en el ámbito de la disponibilidad léxica por primera vez por Ávila Muñoz y Sánchez Sáez (2010, 2011).

6.6.1.1. Léxico según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición

Estos criterios han sido los más usuales en diferentes investigaciones para elegir el léxico más apropiado según el propósito. Samper Padilla (1999) selecciona una muestra representativa del léxico en lugar de trabajar con todas las palabras argumentando que «a partir de un determinado punto en los listados de disponibilidad aparece un conjunto de palabras que han sido aportadas por muy pocos encuestados, que, además, las han incluido en sus hojas de respuestas en posiciones poco relevantes».

Las primeras tres columnas se refieren al índice de disponibilidad calculado por el programa Lexidisp. Ruiz Basto (1987: 97) establece la clasificación entre

vocabulario activo y pasivo dividiendo el mismo entre vocablos con un índice de disponibilidad superior a 0,10 e inferior. Además, añade que para que el vocabulario se considere activo, hace falta que el mismo sea mencionado por al menos un 20 % de los informantes. De esta forma, en su investigación del léxico en centros escolares mexicanos esa clasificación la realiza tomando dos fuentes diferentes: el índice de disponibilidad y el porcentaje de aparición.

Este criterio ha sido utilizado y modificado en posteriores estudios. Así, Carcedo González (2000c: 157-211) compara el léxico con un índice de disponibilidad de al menos 0,1 de cuatro comunidades hispanohablantes y de estudiantes finlandeses de español. Más tarde, el estudioso asturiano (Carcedo González, 2001: 62) selecciona el vocabulario aportado en Asturias con un índice de 0,1 y un porcentaje de aparición del 30 %.

Este rango de vocablos ha sido más laxo en investigaciones posteriores. Gómez Devís (2004: 125-127) concluye que ante la disparidad de los resultados obtenidos según los criterios de frecuencia acumulada y el índice de disponibilidad léxica mencionado anteriormente, incluye las palabras que superan el 0,05. Bellón Fernández (2011: 182-283) comenta las ventajas de establecer el rango en el índice de disponibilidad de $\geq 0,05$. Por una parte, considera que «el número de términos que limita es muy uniforme y muy estable para los listados de léxico disponible de los diferentes centros de interés». Por otra parte, menciona que las palabras resultantes suelen estar en torno a las 50 lexías idóneas propuestas por Alba Ovalle (1998, 2000b). Además, esas palabras suelen haber sido producidas por al menos un 10 % de los hablantes.

En esta misma línea, Bartol Hernández (2001: 227-230) amplía la muestra estableciendo como premisa que el índice de disponibilidad fuese igual o superior a 0,02 y hubiese sido mencionado por más de un informante. En la siguiente columna se han incluido los resultados con un índice $\geq 0,01$ por considerar que amplía el vocabulario eliminando las asociaciones personales. Cabe mencionar que ese índice podrá incluir palabras aportadas por un único informante pero se situarán en una de las primeras posiciones en las listas, con lo que la relación con el centro de interés suele ser directa.

La quinta columna muestra los resultados de los vocablos con un índice acumulado del 75 %. Samper Padilla (1999) estableció este criterio para comparar el léxico de Puerto Rico y Gran Canaria. Consiste en ordenar las palabras según un orden ascendente de frecuencia y contabilizar la frecuencia acumulada. Como se puede apreciar en los resultados, es un índice bastante restrictivo por lo que algunos autores como Bartol Hernández (2000, 2001) han cuestionado su relevancia.

En los siguientes cuadros se ha incluido el número de vocablos según los diferentes criterios. El hecho de que se hayan presentado en tres idiomas y con un número de informantes variable permite detectar qué factores influyen en esta selección léxica. Se ha incluido en la última fila el porcentaje de vocablos que representa con respecto al total.

De forma general, se observa que el porcentaje de vocablos incluidos en cada clasificación varía de forma considerable, lo que daría lugar a comparaciones muy diferentes según el criterio empleado. El apartado que cuenta con menos palabras es el de frecuencia acumulada mientras que los criterios del índice de disponibilidad más amplios incluyen gran cantidad de vocablos.

En español, el número de unidades léxicas extraídas según el 75 % de frecuencia acumulada solo cuenta con 90 vocablos de los 3579 que los informantes han actualizado, lo que representa un total de 2,51 %. En el lado opuesto, estarían los vocablos según el índice de frecuencia. En el caso del índice $\geq 0,01$ asciende a 39,15 % y en el del índice $\geq 0,10$ se reduce a 8,27 %. Si además, se considera el criterio del 30 % de los informantes, solo se incluiría el 4,30 % de los vocablos.

CI	ID $\geq 0,10$	ID $\geq 0,05$	ID $\geq 0,02$	ID $\geq 0,01$	ID $\geq 0,10$ y 30 % informantes	Frecuencia Acumulada 75 %	Total
C1	38	49	88	124	28	9	244
C2	37	52	82	105	25	7	238
C3	43	78	140	214	19	15	530
C4	32	56	85	136	21	8	392
C5	36	67	113	182	13	11	483
C6	22	39	63	90	14	5	190
C7	29	58	116	180	12	11	469
C8	34	68	132	199	11	13	478
C9	25	60	109	171	11	11	555
Total	296	527	928	1401	154	90	3579

Media	32,89	58,56	103,11	155,67	17,11	10	397,67
	8,27 %	14,72 %	25,93 %	39,15 %	4,30 %	2,51 %	100 %

Tabla 6.29. Vocablos en español según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición

En inglés, los datos referidos a los vocablos es inferior no solo por contar con 21 informantes menos, sino por tratarse de una lengua extranjera. No obstante, el porcentaje en cada apartado es muy similar al español. Los centros más compactos aportan menor número de palabras en total y en los listados según diversos criterios.

CI	ID 0,10	≥ ID 0,05	≥ ID 0,02	≥ ID 0,01	≥ ID ≥ 0,10 y 30 % informantes	Frecuencia Acumulada 75 %	Total
C1	38	49	88	124	28	9	244
C2	24	32	46	67	16	5	151
C3	34	56	87	121	17	10	287
C4	19	31	55	97	4	5	244
C5	21	52	82	122	9	8	343
C6	18	27	41	60	11	4	165
C7	18	39	82	131	7	7	310
C8	19	42	74	117	6	7	280
C9	24	41	85	143	7	10	460
Total	202	355	600	917	99	62	2355
Media	22,44	39,44	66,67	101,89	11	6,89	261,67
Porcentaje	8,58 %	15,07 %	25,48 %	38,94 %	4,20 %	2,63 %	100 %

Tabla 6.30. Vocablos en inglés según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición

Los pruebas en francés han sido realizadas por 21 informantes lo que influye en todos los resultados. Así, se observa que el porcentaje de vocablos aportados por el 30 % de los informantes asciende a 11,35 % mientras que en español a 4,30 % y en inglés a 4,20 %. Lo mismo ocurre con la diferencia porcentual marcada por el índice de disponibilidad, oscilando entre 23,09 % y el 96,29 %. En los casos con pocos informantes, muchos vocablos han sido aportados por un único individuo. De hecho, el francés cuenta en el apartado anterior con el porcentaje más amplio en este sentido. Por tanto, si se mantiene ese criterio, el número de vocablos se reduciría considerablemente.

CI	ID 0,10	≥ ID 0,05	≥ ID 0,02	≥ ID 0,01	≥ ID ≥ 0,10 y 30 % informantes	Frecuencia Acumulada 75 %	Total
C1	28	40	59	73	25	54	78
C2	32	45	70	80	16	53	80

C3	32	59	122	156	14	93	166
C4	21	36	84	107	9	71	113
C5	33	57	110	129	14	79	135
C6	16	24	42	53	12	40	53
C7	21	45	103	123	7	69	125
C8	23	42	96	116	9	70	119
C9	24	47	92	122	7	75	127
Total	230	395	778	959	113	604	996
Media	25,56	43,89	86,44	106,56	12,56	67,11	110,67
Porcentaje	23,09 %	39,66 %	78,11 %	96,29 %	11,35 %	60,64 %	100 %

Tabla 6.31. Vocablos en francés según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición

6.6.1.2. Determinación del léxico relevante

Uno de los factores clave en la revisión del modelo matemático, comentado en el apartado, se basa en la determinación del léxico relevante. En la Universidad de Málaga Ávila Muñoz y Sánchez Sáez (2010, 2011) y Ávila Muñoz y Villena Ponsoda (2010, 2011) aplicaron la teoría de conjuntos difusos, *Fuzzy Expected Value* (FEV), a los estudios de disponibilidad léxica³⁰⁷. A grandes rasgos, su cálculo permite diferenciar en varios niveles la compatibilidad del léxico con respecto al centro de interés. Cada nivel recoge un número reducido de vocablos que va aumentando según sea menor el grado de compatibilidad. Además, esos parámetros permiten identificar claramente cuáles son que cuentan con una alta disponibilidad en los informantes y, por tanto, que no deberían faltar en un diccionario.

Este FEV es capaz de proponer un valor de pertenencia que establece un equilibrio entre el número de términos que lo supera y el propio valor. Además, si ponderamos en diferente grado la relevancia del tamaño del conjunto de corte, se pueden establecer grados de restricción en su interior que nos permitirían construir conceptos como «muy representativo» o «poco representativo» (Ávila Muñoz y Villena Ponsoda, 2010: 63).

Los resultados obtenidos permiten diferenciar hasta seis niveles de compatibilidad con el área temática. Cuanto más se acerquen los valores a 1, mayor será su representatividad en el área temática. El corte en niveles en cada centro varía en función de cada estudio y cada centro y vendrá determinado por el FEV.

³⁰⁷ Véase Ávila Muñoz y Villena Ponsoda (2010: 55-65), Ávila Muñoz y Sánchez Sáez (2011), Ávila Muñoz y Villena Ponsoda (2012), Ávila Muñoz y Sánchez Sáez (2014: 260-264).

En lengua española, de los 244 vocablos actualizados en el primer centro de interés, solo 43 forman parte de los cinco niveles considerados centrales, es decir, un 17,62 %. En ese mismo centro, en inglés han sido seleccionados 26 vocablos de las 115 aportadas, correspondiendo a un 22,61 % del total. En francés de las 78 lexías diferentes se han extraído 22, lo que supone un 28,21 %.

Estos datos son de gran utilidad en la selección del léxico como se abordará en los apartados 6.6.3 y 6.6.4, sobre todo a nivel pedagógico, ya que estas serían las primeras palabras en relación al cuerpo humano que el estudiante debería aprender. A continuación, se presenta a modo de ejemplo los resultados en los tres idiomas para el primer centro de interés³⁰⁸.

Nivel de compatibilidad	Proporción	Total 224	Centro 1
0,88	1,64%	4	cabeza, ojo, mano, pie
0,82	4,51%	11	dedo, brazo, nariz, boca, oreja, pierna, pelo
0,69	4,92 %	12	uña
0,51	6,97 %	17	cara, cuello, ceja, hombro, codo
0,33	11,07 %	27	diente, espalda, rodilla, pestaña, pecho, labio, lengua, tobillo, tronco, corazón
0,18	17,62 %	43	muñeca, frente, estómago, mejilla, pulmón, cráneo, barriga, cadera, hueso, culo, antebrazo, cerebro, cintura, muslo, riñón, barbilla

Tabla 6.32. Léxico nuclear en español del centro de interés 01. *Partes del cuerpo humano*

Nivel de compatibilidad	Proporción	Total 115	Centro 1
0,91	4,35 %	5	head, eye, arm, leg, hand
0,84	6,09 %	7	nose, mouth
0,74	8,7 %	10	finger, ear, hair
0,58	11,3 %	13	shoulder, foot, neck
0,43	18,26 %	21	face, feet, knee, teeth, nail, toe, heart, elbow
0,23	22,61 %	26	lip, tongue, back, eyebrow, chest

Tabla 6.33. Léxico nuclear en inglés del centro de interés 01. *Human body parts*

Nivel de compatibilidad	Proporción	Total 78	Centro 1
0,89	2,56	2	tête, main
0,85	6,41	5	oeil, pied, bras
0,76	10,26	8	nez, oreille, jambe
0,61	14,1	11	bouche, cou, yeux
0,44	19,23	15	doigt, cheveu, cheville, genou
0,28	28,21	22	ongle, dent, ventre, dos, coude, langue, épaule

Tabla 6.34. Léxico nuclear en francés del centro de interés 01. *Parties du corps humain*

³⁰⁸ El resto del léxico compatible se encuentra en el Anexo VIII.

6.6.2. Los aspectos socioculturales en el léxico disponible

La procedencia de un hablante puede hacer que el término *persiana* no aparezca como parte de la casa. La religión puede hacer que *mezquita* o *iglesia* sean los lugares sagrados más disponibles en el centro de interés de *La ciudad* o que el *Corán* o la *Biblia* estén relacionados con *La escuela*. La edad y las costumbres pueden hacer que ciertos alimentos y bebidas alcohólicas estén presentes en el léxico nuclear. En definitiva, el momento concreto y la particularidad de los hablantes pueden hacer del léxico disponible un diccionario sincrónico y representativo de una sociedad.

Tras analizar los diccionarios de léxico disponible en español como lengua materna, español como segunda lengua y español como lengua extranjera se constatan una serie de divergencias que muestran la diversidad de culturas y costumbres en diferentes países. Por ejemplo, en la investigación realizada en Málaga sobre una población adulta estratificada (Ávila Muñoz y Villena Ponsoda, 2010) recogen en el centro de interés denominado *Los muebles de la casa* el término *persiana*, concretamente en la posición 130 de 193 vocablos. Del mismo modo, en las respuestas de inmigrantes en Castilla la Mancha del estudio de Fernández-Merino Gutiérrez (2015)³⁰⁹ aparece este término en el centro *La casa*, concretamente en la posición 85 de 257 vocablos.

Sin embargo, los estudiantes fineses de español de la muestra recogida por Carcedo González (2000c) no citan *persiana* en el centro referido a la casa. Como es bien sabido, en Finlandia no se utilizan las persianas para graduar la entrada de la luz, tan solo en algunas ocasiones se usan las cortinas, término que tampoco recoge el mencionado diccionario de léxico disponible. En el estudio coordinado por Bagri y Samper Padilla (2013), los estudiantes universitarios de español como lengua extranjera en Marruecos no mencionan el término *persiana*, pero *cortina* aparece en el puesto 28 de 64 vocablos según el índice de disponibilidad léxica.

Cabe destacar que la muestra de estudiantes marroquíes (Bagri y Samper Padilla, 2013) no recoge algunos términos muy disponibles y considerados nucleares

³⁰⁹ Corresponde a la tesis doctoral presentada en el año 2013 en la Universidad de Valladolid. Disponible en: <http://uvadoc.uva.es/bitstream/10324/4373/1/TESIS464-140219.pdf> [Fecha de última consulta el 5 de marzo de 2015]

en el presente estudio y, por tanto, imprescindibles en cualquier manual de español para extranjeros. En el centro de interés *Alimentos y bebidas* los informantes de español para extranjero no aportan ningún vocablo que haga referencia a una bebida alcohólica mientras que en los estudiantes de postgrado se registra el vocablo *cerveza* como el segundo más compatible. Del mismo modo, tampoco está presente *iglesia* en el centro denominado *La ciudad*, vocablo que forma parte del sexto nivel de compatibilidad en esta la muestra actual.

Dado que la mayoría de la población marroquí es musulmana, en torno al 98 %, la mezquita es considerada uno de los edificios más importantes de la comunidad. En Bagri y Samper Padilla (2013) *mezquita* aparece dentro de los veinte primeros vocablos tanto de la muestra de estudiantes de nivel inicial como de universitarios. *Mezquita* también está presente en los estudios de español como lengua materna (Ávila Muñoz y Villena Ponsoda, 2010) y de español como segunda lengua (Fernández-Merino Gutiérrez, 2015). En cambio, no aparece en el léxico de español como lengua extranjera de los estudiantes finlandeses (Carcedo González, 2000c), país donde el número de mezquitas es muy reducido y el número de musulmanes ronda el 0,2 %.

Ese tipo de comparaciones se puede realizar en todos los centros de interés y muestran la forma de vestir de los informantes (*chilaba*), sus comidas típicas (*cuscús*, *tajín*, *ejote*), un uso diferente de las nuevas tecnologías (*ciber*), la decoración (*alfombra*, *tapiz*), el diseño de la ciudad (*zoco*), las ocupaciones (*religioso*), etc. En otras ocasiones, las palabras muestran una realidad mucho más compleja como la forma de vida (bastantes palabras en el centro de interés del campo), el clima (*agua fría*, *bebida fresca*), la forma de comprar la comida (*matar* aparece en el centro de los alimentos porque en algunos establecimientos pesan y matan al animal, normalmente un pollo, en el momento de comprarlo), la doctrina religiosa (*cerdo* no aparece ni en el centro relacionado con la comida ni con los animales).

Los diccionarios dan cuenta también del diacronismo de la lengua. Por ejemplo, Alba Ovalle (2011) realiza un estudio comparativo de dos muestras de la República Dominicana con una diferencia de casi dos décadas (los datos del primer estudio fueron recogidos en 1989 y los del segundo en 2008). En el análisis

cualitativo, se destaca que algunos vocablos mantienen su disponibilidad léxica, como es el caso de *cerdo* en el centro denominado *Animales* mientras que otros pierden disponibilidad, como por ejemplo *puerco*. Ese cambio supone que el léxico disponible de los dominicanos se acerca a la variedad peninsular donde la palabra *cerdo* es considerablemente más disponible que *puerco* ya que en ocasiones ni siquiera aparece en los listados, como ocurre en el centro referido a los animales del estudio de Ávila Muñoz y Villena Ponsoda (2010).

Además de los factores socioculturales, la edad de los informantes también influye en el léxico y permite conocer y analizar el comportamiento de una generación a través de las respuestas de cada centro de interés. No se vestirá igual un adolescente que un adulto por lo que vocablos como *minifalda*, *leggings*, *top* y *bailarinas* tendrán mayor disponibilidad entre las chicas jóvenes. Tampoco será igual su forma de divertirse (*chatear*, *whasappear*, *botellón*, *salir de marcha*, *patinar*, etc.), de comer y beber (*pizca*, *espaguetis*, *hamburguesa*, *ron*, *ginebra*) o de expresar su relación con respecto a la escuela (*aburrimiento*, *examen*, *hacer novillos*) o las nuevas tecnologías (*Facebook*, *Twitter*, *selfie*). Samper Hernández (2003: 977) indica que los estudiantes preuniversitarios de Gran Canaria incorporan en sus listados de léxico disponible nombres de bebidas alcohólicas con una mayor disponibilidad que los estudiantes extranjeros debido a la costumbre española de los jóvenes y quizás a la mayor edad de los informantes de los extranjeros.

Llegados a este punto, ¿qué léxico se debería enseñar? ¿tiene el léxico que mostrar la cultura de la comunidad de hablantes de la lengua de estudio o de la lengua materna? En el siguiente apartado se abordará la selección del léxico según su compatibilidad. Obviamente, se incluyen palabras que ayudan a describir el entorno extralingüístico de los informantes y que son parte integrante de la cultura. No obstante, no se puede olvidar que el contexto de aprendizaje de la lengua influye en la selección del léxico eliminando palabras que carezcan de interés e incluyendo otras que sean determinantes. Por ello, el docente ha de saber que no sería recomendable que una de las primeras palabras relacionadas con la bebida que aprendiesen los escolares menores fuese *cerveza* y, más aún en países donde las bebidas alcohólicas están más restringidas.

6.6.3. El léxico nuclear y el nivel según el Marco Común Europeo de Referencia para las Lenguas

El Consejo de Europa llevó a cabo durante el año 2001 un impulso reseñable en relación al aprendizaje de las lenguas mediante la celebración del «Año europeo de las lenguas» y la publicación del *Marco Común Europeo de Referencia para las Lenguas*. De esta forma, todos los estados pertenecientes se centraron en desarrollar la comunicación y la interacción entre sus ciudadanos, con el fin de fomentar la movilidad, la colaboración y la comprensión mutua, insistiendo en la necesidad del aprendizaje de las lenguas como medio para conseguir estos propósitos.

El Marco Común Europeo de Referencia para las Lenguas tiene como objeto servir de guía a los profesionales del ámbito de la enseñanza de idiomas para diseñar programas de estudio, elaborar materiales curriculares y establecer pautas universales para la certificación y evaluación de competencias lingüísticas en Europa. Bartol Hernández (2010: 94-95) considera que existe una gran similitud entre los centros de interés tradicionales y los temas y subtemas que desarrollan los ámbitos del Consejo de Europa. A través de la Tabla 6.35 el investigador muestra la correspondencia, más o menos parcial:³¹⁰

Centros de interés (Disponibilidad léxica)	Temas y subtemas del Consejo de Europa
01. Partes del Cuerpo	7.1. Partes del cuerpo
02. La ropa	9.3. Ropa y moda
03. Partes de la casa	2. Vivienda
04. Los muebles de la casa	2. Vivienda
05. Comidas y bebidas	10. Comidas y bebidas
06. Objetos colocados en la mesa para la comida	
07. La cocina y sus utensilios	2. Vivienda
08. La escuela	8. Educación ¿?
09. Iluminación, calefacción y medios de airear un recinto	
10. La ciudad	9. Compras
	11. Servicios públicos
	12. Lugares
11. El campo	12. Lugares
	2.8. Flora y fauna
12. Medios de transporte	5.1. Transporte público
	5.2. Transporte privado
13. Trabajos del campo y el jardín	
14. Animales	2.8. Flora y fauna

³¹⁰ No obstante, Bartol Hernández (2010: 104) afirma que los centros de interés tradicionales no abarcan temas imprescindibles y aboga por una revisión de los mismos.

15. Juegos y diversiones	4. Tiempo libre, entretenimiento
16. Profesiones y oficios	1.10. Profesión, trabajo

Tabla 6.35. Tabla comparativa de los centros de interés y los temas y subtemas del Consejo de Europa, en Bartol Hernández (2010: 95)³¹¹

Siguiendo las recomendaciones de este marco de referencia, en el año 2006 se publica el Plan Curricular del Instituto Cervantes (PCIC) para concretar y desarrollar los niveles en español. Esta obra se divide en tres tomos que corresponde a los niveles del usuario básico (A1 y A2), del usuario independiente (B1 y B2) y del usuario competente (C1 y C2). El PCIC incluye dos inventarios nocionales que recogen las especificaciones de contenidos relacionados con el significado basados en el enfoque nociofuncional de los años 70. Los inventarios no solo incorporan palabras sino también unidades léxicas pluriverbales, colocaciones y expresiones idiomáticas.

El primer inventario, denominado *Nociones Generales*, se refiere a las unidades léxicas predominantemente abstractas que un hablante puede utilizar en cualquier situación comunicativa. Se divide en las siguientes siete categorías que a su vez están compuestas de subcategorías:

1. Nociones existenciales
2. Nociones cuantitativas
3. Nociones espaciales
4. Nociones temporales
5. Nociones cualitativas

El segundo inventario, *Nociones específicas*, se refiere a situaciones concretas relacionadas con un tema determinado. La agrupación de los centros de interés en campos nocionales hace posible la comparación con el inventario de nociones específicas. A continuación, se ha establecido un relación entre las nociones específicas presentes en el PCIC y los centros de interés tradicionales.

³¹¹ Bartol Hernández (2010: 95) hace a su vez referencia al trabajo de Ángela Bombarelli (2005).

Nociones específicas del Plan Curricular del Instituto Cervantes		Centros de interés tradicionales
1. Individuo: dimensión física	1.1. Partes del cuerpo	1. Partes del cuerpo
	1.2. Características físicas	
	1.3. Acciones y posiciones que se realizan con el cuerpo	
	1.4. Ciclo de la vida y reproducción	
2. Individuo: dimensión perceptiva y anímica	2.1. Carácter y personalidad	
	2.2. Sentimientos y estados de ánimo	
	2.3. Sensaciones y percepciones físicas	
	2.4. Estados mentales	
	2.5. Modales y comportamiento	
	2.6. Valores personales	
	2.7. Suerte	
3. Identidad personal	3.1. Datos personales	3.1.1. Nombre
		3.1.2. Dirección
		3.1.3. Número de teléfono
		3.1.4. Lugar y fecha de nacimiento
		3.1.5. Nacionalidad
		3.1.6. Edad
		3.1.7. Sexo
		3.1.8. Estado civil
		3.1.9. Profesión
	3.2. Documentación	16. Profesiones y oficios
4. Relaciones personales	4.1. Relaciones familiares	
	4.2. Relaciones sociales	
	4.3. Celebraciones y actos familiares, sociales y religiosos	
	4.4. Actitudes y formas de comportarse	

5. Alimentación	5.1. Dieta y nutrición	
	5.2. Bebida	5. Comidas y bebidas
	5.3. Alimentos	
	5.4. Recetas	
	5.5. Platos	
	5.6. Utensilios de cocina y mesa	6. Objetos colocados en la mesa para la comida ²
	5.7. Restaurante	7. La cocina y sus utensilios
6. Educación	6.1. Centros e instituciones educativas	
	6.2. Profesorado y alumnado	
	6.3. Sistema educativo	
	6.4. Aprendizaje y enseñanza	
	6.5. Exámenes y calificaciones	
	6.6. Estudios y titulaciones	
	6.7. Lenguaje de aula	
	6.8. Material educativo y mobiliario de aula	8. La escuela: muebles y materiales
7. Trabajo	7.1. Profesiones y cargos	16. Profesiones y oficios
	7.2. Lugares, herramientas y ropa de trabajo	
	7.3. Actividad laboral	
	7.4. Desempleo y búsqueda de trabajo	
	7.5. Derechos y obligaciones laborales	
	7.6. Características de un trabajador	
8. Ocio	8.1. Tiempo libre y entretenimiento	15. Juegos y distracciones
	8.2. Espectáculos y exposiciones	
	8.3. Deportes	
	8.4. Juegos	

³¹² *Platos* se refiere al tipo de alimento no al recipiente. En cambio, el apartado 5.6. *Utensilios de cocina y mesa* comprende dos centros de interés diferentes, tanto los objetos que se colocan en la mesa como los utilizados para cocinar.

9. Información y medios de comunicación	9.1. Información y comunicación	
	9.2. Correspondencia escrita	
	9.3. Teléfono	
	9.4. Prensa escrita	
	9.5. Televisión y radio	
	9.6. Internet	
10. Vivienda	10.1. Acciones relacionadas con la vivienda	10.1.1. Construcción
		10.1.2. Compra y alquiler
		10.1.3. Ocupación
	10.2. Características de la vivienda	10.2.1. Tipos
		10.2.2. Partes
		3. Partes de la casa
		10.2.3. Personas
		10.2.4. Condiciones
		9. Iluminación, calefacción y medios de airear un recinto
	10.3. Actividades domésticas	10.3.1. Limpieza de la casa
		10.3.2. Decoración de la casa
11. Servicios	10.4. Objetos domésticos	10.4.1. Muebles y objetos domésticos
		4. Muebles de la casa
	10.4.2. Electrodomésticos	
	11.1. Servicio postal	
	11.2. Servicios de transporte ³	12. Medios de transporte
	11.3. Servicios financieros	
	11.4. Servicios sanitarios	
	11.5. Servicios educativos	
	11.6. Servicios de protección y seguridad	
	11.7. Servicios sociales	
12. Compras,	11.8. Servicios de abastecimiento público	
	12.1. Lugares, personas y actividades	

³¹³ El PCIC en el apartado *Servicios de transporte* en todos los niveles redirige a Nociones específicas 14.3. *Sistemas de transporte*.

tiendas y establecimientos	12.2. Ropa, calzado y complementos	2. La ropa	
	12.3. Alimentación		
	12.4. Pagos		
13. Salud e higiene	13.1. Salud y enfermedades		
	13.2. Heridas y traumatismos		
	13.3. Síntomas		
	13.4. Centros de asistencia sanitaria		
	13.5. Medicina y medicamentos		
	13.6. Higiene		
	13.7. Estética		
14. Viajes, alojamiento y transporte	14.1. Viajes	14.1.1. Objetos y documentos relacionados con los viajes	
		14.1.2. Tipos de viajes	
	14.2. Alojamiento		
	14.3. Sistema de transporte	14.3.1. Red de transportes	12. Medios de transporte
		14.3.2. Tipos de transporte	
		14.3.3. La conducción	
15. Economía e industria	15.1. Finanzas y bolsa	15.1.1. Economía y dinero	
		15.1.2. Mercado financiero	
		15.1.3. Organismos e instituciones financieros y mercantiles	
	15.2. Renta		
	15.3. Comercio	15.3.1. Transacciones comerciales y mercados	
		15.3.2. Comercio exterior	
		15.3.3. Publicidad, mercadotecnia	
	15.4. Entidades y empresas	15.4.1. Tipos de empresa y organización	
		15.4.2. Situación de la empresa	
	15.5. Industria y energía	15.5.1. Construcción, industria pesada y ligera	
		15.5.2. Sector agropecuario	13. Trabajos del campo y del jardín
			14. Los animales
		15.5.3. Pesca	

16. Ciencia y tecnología	16.1. Cuestiones generales		
	16.2. Biología		
	16.3. Matemáticas		
	16.4. Informática y nuevas tecnologías		
	16.5. Física y química		
17. Gobierno, política y sociedad	17.1. Sociedad	17.1.1. Vida en comunidad	
		17.1.2. Conducta social	
	17.2. Política y gobierno	17.2.1. Instituciones políticas y órganos de gobierno	
	17.3. Ley y justicia		
	17.4. Ejército		
18. Actividades artísticas	18.1. Disciplinas y cualidades artísticas		
	18.2. Música y danza		
	18.3. Arquitectura, escultura y pintura		
	18.4. Literatura		
	18.5. Fotografía		
	18.6. Cine y teatro		
19. Religión y filosofía	19.1. Religión		
	19.2. Filosofía		
20. Geografía y naturaleza	20.1. Universo y espacio		
	20.2. Geografía	20.2.1. Geografía física, humana y política	
		20.2.2. Paisaje y accidentes geográficos	
	20.3. Espacios urbanos o rústicos	20.3.1. Ciudad	10. La ciudad
		20.3.2. Campo	11. El campo 13. Trabajos del campo y del jardín
	20.4. Clima y tiempo atmosférico		
	20.5. Fauna		14. Los animales
	20.6. Flora		13. Trabajos del campo y del jardín
	20.7. Problemas medioambientales y desastres naturales		

Tabla 6.36. Relación entre las nociones específicas y los centros de interés tradicionales

A diferencia de la Tabla 6.35, la relación se ha realizado contrastando los temas del PCIC con respecto a los centros de interés para conocer el alcance semántico del léxico disponible. El cotejo de los campos nocionales muestra que los centros tradicionales no abarcan ni la mitad de los ámbitos propuestos por el PCIC. Asimismo, en ciertas ocasiones la correspondencia no es directa y es frecuente que el centro de interés haga referencia a un subtema sin llegar a abarcar el ámbito concreto del PCIC.

A modo de ejemplo, el centro de interés *Iluminación, calefacción y medios de airear un recinto* se encuentra presente en diversos apartados:

- 10.2.4. Condiciones. Nivel A1: *tener aire acondicionado/calefacción (central/individual)*.
- 14.2. Alojamiento. Nivel A1: *habitación con aire acondicionado*.
- 10.3.2. Decoración de la casa. Nivel B1: *iluminar una habitación/un espacio*. Nivel B2: *iluminación/luz ~ directa/indirecta/suave/tenue/alegre*.
- 14.4.2. Electrodomésticos. Nivel B2: *bombilla*.

Del mismo modo, el centro de interés *Trabajos del campo y del jardín* aparece disperso en varios temas:

- 20.6. Flora. Nivel B1: *plantar un árbol, regar las plantas/el jardín, cultivar, cortar una flor*.
- 20.3.2. Campo. Nivel B2: *trasplantar, talar un árbol, deshojar una flor*. Nivel C2: *arar, labrar, segar, recolectar, siembra, labranza, matanza*. Ambos niveles, reflejan que las labores del campo no forman parte de un vocabulario básico ya que aparecen en el usuario independiente y competente.
- 15.5.2. Sector agropecuario. Nivel C1: *cosechar, ordeñar*. Nivel C2: *podar, trillar, esquilar, labrar, fermentar*.

La relación de los otros centros mantiene una correspondencia más directa con los temas o subtemas del PCIC. El centro *Profesiones y Oficios* aparece en 3.1.9. Profesión y 7.1. Profesiones y cargos; *Animales* en 20.5. Fauna y 15.5.2. Sector agropecuario y el resto de centros de interés podrían distribuirse en los temas o subtemas asignados. Cabe destacar que para simplificar el parangón, se ha intentado realizar la asociación con el tema o centro al que pertenece por antonomasia

obviando las particularidades de algunas unidades léxicas presentes en un centro y no en el tema y viceversa.

Según Paredes García (2014) precisamente los dos centros de interés más dispersos en el PCIC *Trabajos del campo y del jardín* e *Iluminación, calefacción y medios de airear un recinto* «parecen candidatos idóneos para ser excluidos de las encuestas» dada su poca productividad. También se ha criticado que han perdido importancia hoy en día y que están desfasados, razón por la cual no suponen una categoría nocional tan bien definida como el resto y no se enseña en los niveles básicos.

González Fernández (2014) evalúa la idoneidad de los centros de interés clásicos en una investigación en Turquía sobre el léxico disponible en español como lengua extranjera. A través de un análisis cuantitativo sobre la producción léxica en cada centro de interés, deja entrever cuáles son los centros que deberían ser modificados por otros más productivos. Concretamente se refiere a los centros con menor número de palabras en estudios previos (*Trabajos del campo y el jardín*, *El campo*, *La calefacción y la iluminación* y *Los objetos colocados a la mesa para la comida*) y los centros que más de un 30 % de los informantes han dejado en blanco (*El campo*, *Los trabajos del campo y el jardín*, *La ciudad*, *Los juegos y las distracciones*, *La calefacción y la iluminación* y *Los objetos colocados a la mesa para la comida*).

En este sentido son muchos los autores Bombarelli (2005), Bartol Hernández (2010), Sánchez-Saus Laserna (2011) Paredes García (2014) González Fernández (2014) que abogan por una renovación de los centros de interés. Defienden que si bien es cierto que los centros clásicos permiten la comparación con estudios precedentes, esos centros no se ajustan a las nociones básicas referentes a la vida cotidiana presentes en el MCERL.

6.6.3.1. Léxico nuclear en los idiomas del estudio

El léxico nuclear o más compatible varía en función de las respuestas de cada muestra. En este estudio se presenta en el Anexo VII el listado de palabras más compatibles en los tres idiomas. El listado en español es fruto del uso de la lengua en un contexto natural mientras que en la mayoría de las ocasiones el léxico aportado en inglés y francés ha sido adquirido y utilizado en exclusiva en el ámbito académico.

En consonancia con las reflexiones sobre el comportamiento de los informantes de diferentes culturas en las pruebas de disponibilidad léxica, se presentan los listados con el léxico nuclear del centro de interés *Comidas y bebidas* en las tres lenguas para comparar los resultados. Su selección se debe al marcado carácter sociocultural de este ámbito con respecto a otros centros de interés como puede ser *Partes del cuerpo humano*. De hecho, el inventario denominado *Saberes y comportamientos socioculturales* del PCIC dedica un apartado a las comidas y bebidas destacando no solo las unidades léxicas concretas que lleva este centro asociadas (*tapa*) sino también los conocimientos específicos relativos al comportamiento compartido por los hablantes. Por ejemplo, qué es la merienda y cuándo se toma, cuál es el horario de cada comida, cómo se organiza el menú y, en definitiva, cuáles son las convenciones sociales y los comportamientos en la mesa.

En el ámbito de los estudios hispánicos en Serbia y la región, Pejović, Sekulić y Karanović editan el volumen monográfico *Comida y bebida en la lengua española, cultura y literaturas hispánicas* (2012). Se trata de un compendio de investigaciones que analizan el tema de la comida y la bebida desde un enfoque interdisciplinar en el que interviene además de la lengua, la literatura y la cultura de los países hispanos.

Šifrar Kalan (2012: 191-208) dedica un artículo a mostrar la relación entre el léxico disponible y la cultura a través de un análisis comparativo de los listados de estudiantes eslovenos de español como lengua extranjera y de alumnos preuniversitarios de Asturias, Cádiz y Valencia. Asimismo, contrasta los resultados de los estudiantes españoles con el vocabulario presente en dos manuales de español como lengua extranjera (Gente 1 y Gente 2) y el PCIC. El estudio concluye que existe una estrecha relación entre el léxico de los estudiantes de lengua materna y lengua extranjera y que el léxico disponible de los preuniversitarios está presente tanto en los manuales como en el PCIC.

Dadas las características particulares de este estudio, el análisis se va a realizar cotejando los resultados en los tres idiomas para comprobar si el alumno tiende a repetir las mismas palabras o se ve influido por otros contextos de aprendizaje, como puede ser el manual. *A priori*, llama la atención que el número de palabras compatibles en español sea superior al inglés y el francés. Este hecho se produce

como consecuencia del mayor número de informantes (ya que todos han contestado en español) y de la productividad de los centros de interés en lengua materna. Por contraposición, la proporción de los vocablos con respecto al conjunto es inferior en español así como el nivel de compatibilidad que delimita el léxico más compatible.

Nivel de compatibilidad	Proporción	Total 531	Centro 3
0,84	0,19 %	1	agua
0,76	1,13 %	6	cerveza, carne, pan, Coca-Cola, zumo
0,62	2,07 %	11	leche, pescado, patata, vino, café
0,46	4,33 %	23	tomate, pasta, arroz, lechuga, paella, refresco, fruta, huevo, ensalada, lenteja, chocolate, verdura
0,28	7,35 %	39	manzana, tortilla, naranja, té, sopa, pollo, pizza, hamburguesa, potaje, filete, puchero, Fanta, espagueti, tortilla de patatas, bocadillo, plátano
0,14	13,74 %	73	macarrón, batido, zanahoria, jamón, galleta, ternera, limón, cebolla, güisqui, tarta, queso, fresa, legumbre, melón, atún, ron, pimienta, cereal, pera, aceite, chorizo, cerdo, sal, cocido, gazpacho, salchicha, calabaza, maíz, sandía, mantequilla, garbanzo, solomillo, azúcar, yogur

Tabla 6.37. Léxico nuclear en español del centro de interés 03. *Comidas y bebidas*

El listado en español comprende palabras que no están presentes en inglés y francés y viceversa. En inglés aparecen cinco términos que no se encuentran en español: *chip* que hace referencia a las patatas fritas; *sandwich* podría considerarse que se encuentra en su forma española (*bocadillo*) aunque se ha marcado porque en español puede hacer referencia al tipo de bocadillo que se prepara con pan de molde; *bean* que parece tener mayor relevancia en la dieta anglosajona; *pineapple* puede ser más frecuente porque su venta suele ser en almíbar e incluye la traducción en varios idiomas (a diferencia de otras frutas y verduras) y *pepper* suele enseñarse al mismo tiempo que *salt* aunque no su uso no sea tan frecuente, razón por la cual en español no aparece en el listado.

Nivel de compatibilidad	Proporción	Total 287	Centro 3
0,86	0,7 %	2	water, meat
0,77	1,39 %	4	Coke, fish
0,65	2,79 %	8	wine, potato, tomato, beer
0,51	6,62 %	19	milk, orange, hamburger, apple, pizza, chip , bread, vegetable, egg, rice, juice
0,34	11,15 %	32	coffee, banana, carrot, soup, chicken, cake, Coca-Cola, fruit, lemon, tea, salad, lettuce, onion
0,18	17,42 %	50	spaghetti, sandwich , beef, cheese, chocolate, bean , pasta, watermelon, sugar, pineapple , omelette, steak, whisky, strawberry, salt, pepper , soda, butter

Tabla 6.38. Léxico nuclear en inglés del centro de interés 03. *Food and drink*

En francés, a pesar de la mayor similitud lingüística, existen 10 vocablos que no están incluidos en el listado en español. Cinco vocablos son de procedencia francesa: *champagne*, *crêpe* y *croissant* son tres palabras típicas de la gastronomía francesa que forman parte del léxico disponible español con una mejor compatibilidad; *chocroute* no aparece en español; *confiture* aparece como *mermelada* pero podría aparecer también como *confitura*. Al igual que en inglés se incluye *frite*, que quiere decir *patata frita* (aunque en francés *chip* hace referencia a las patatas fritas envasadas o de bolsa y *frite* a las patatas fritas tiernas y normalmente alargadas. Las otras tres palabras a pesar de aparecer en el léxico disponible en español no muestran la compatibilidad necesaria para ser seleccionadas en el léxico nuclear: *citronnade*, *dîner* y *canard*.

En el ámbito de la enseñanza de español o francés como segunda lengua o lengua extranjera, resultaría interesante enseñar el vocablo *pato* y *canard* incluso a alumnos principiantes por dos razones: por ser más frecuente su consumo en Francia y por tratarse de un falso amigo. Si en el léxico disponible no se enseña *canard* como *pato*, el estudiante español cuando viaje a puede pedir un plato de *pato* creyendo que está pidiendo *carne* y viceversa.

Nivel de compatibilidad	Proporción	Total 166	Centro 3
0,85	0,6 %	1	eau
0,79	1,81 %	3	fromage, chocolat
0,65	3,01 %	5	omelette, bière
0,53	7,83 %	13	jambon, lait, croissant , poisson, vin, pâte, pain, frite
0,4	15,66 %	26	boeuf, oignon, canard , pomme, viande, thé, choucroute , citronnade , dîner , fruit, pomme de terre, orange, Coca-Cola
0,23	22,29 %	37	vin blanc , pizza, crêpe , café, confiture , champagne , beurre, oeuf, poulet, légume, salade

Tabla 6.39. Léxico nuclear en francés del centro de interés 03. *La nourriture et les boissons*

6.6.4. Léxico nuclear en el Plan Curricular del Instituto Cervantes

Las nociones específicas del Plan Curricular del Instituto Cervantes incorporan un tema denominado *Alimentación* que se subdivide a su vez en siete apartados: dieta y nutrición; bebida; alimentos; recetas; platos; utensilios de cocina y mesa; y restaurante. Los dos subcampos relacionados con el centro de interés *Comidas y bebidas* serían el segundo y tercero (bebidas y alimentos) cuyas unidades léxicas, expresiones idiomáticas y colocaciones se presentan según el nivel del MCER. A continuación, se presentan las nociones específicas del nivel A1 y A2 para poder

relacionarlo con lo que se considera vocabulario compatible y, por consiguiente, el léxico necesario para un usuario básico.

5.2. Bebida	
A1	A2
agua (mineral) ~ con gas/sin gas	bebida ~ con/sin ~ alcohol/hielo
leche, té, café ~ solo/con leche/cortado	zumos ~ de naranja/de tomate/de frutas
cerveza, vino ~ blanco/tinto	infusión
beber	ron, ginebra, cava, champán
	botella, lata
	tener sed
	tomar algo

Tabla 6.40. Nociones específicas del nivel A1 y A2 del PCIC. Apartado 5.2. Bebida

5.3. Alimentos	
A1	A2
carne, pescado, fruta, verdura, leche, huevos, pan	lechuga, tomate, zanahoria
bocadillo, sándwich, hamburguesa	plátano, manzana, naranja
	patata, ajo, cebolla, perejil
	queso, yogur, mantequilla, mayonesa, salsa
	merluza, salmón, sardina, atún, gambas
	pasta, arroz
	chocolate, galletas, cereales
	azúcar, sal, pimienta, aceite
	carne de ~ ternera/cerdo/ cordero/pollo
	filete
	jamón ~ serrano/York
	helado de ~ chocolate/fresa/vainilla
	tarta de ~ manzana/crema/chocolate

Tabla 6.41. Nociones específicas del nivel A1 y A2 del PCIC. Apartado 5.3. Alimentos

En el inventario de saberes y conocimientos socioculturales se incluye información adicional que complementa las nociones específicas. Por ejemplo, aparecen algunos platos típicos por zonas o regiones con independencia del nivel del MCER: *gazpacho* (Andalucía), *fabada* (Asturias), *migas* (Castilla-La Mancha), *bacalao al pilpil* (País Vasco), *paella* (Valencia); el nombre de los condimentos, hierbas y especias más utilizadas en la cocina española (*aceite de oliva, sal, vinagre, pimentón, ajo, azafrán y perejil*), los tipos de platos según su forma de cocina (*cocido, guisado, a la plancha, asado, frito*) entre otros temas que ayudan a mejorar la competencia comunicativa del hablante en este tema concreto.

La Tabla 6.42 presenta el léxico más prototípico del centro de interés *Comidas y bebidas* organizado según los seis niveles de compatibilidad. A la derecha se ha incluido el nivel según MCER y el tema específico que asigna el PCIC. Al igual que en el PCIC se han marcado con el color naranja los niveles A1 y A2 y con el color rojo a los niveles B1 y B2. No se ha usado el color verde que identifica a los niveles C1 y C2 porque no hay ninguna palabra que se corresponda con estos niveles.

Nivel	Vocablo	Compatibilidad	Nivel del MCER	Noción específica del PCIC
Nivel 1	agua	0,955196	A1	5.2. Bebida
Nivel 2	carne	0,81709	A1	5.3. Alimentos
	cerveza	0,835115	A1	5.2. Bebida
	Coca-Cola	0,787496		
	pan	0,791448	A1	5.3. Alimentos
	zumos	0,766884	A2	5.2. Bebida
Nivel 3	café	0,631503	A1	5.2. Bebida
	leche	0,736062	A1	5.2. Bebida
	patata	0,682504	A2	5.3. Alimentos
	pescado	0,699956	A1	5.3. Alimentos
	vino	0,675749	A1	5.2. Bebida
Nivel 4	arroz	0,569067	A2	5.3. Alimentos
	chocolate	0,465749	A2	5.3. Alimentos
	ensalada	0,501331	A1	5.5. Platos
	fruta	0,515579	A1	5.3. Alimentos
	huevo	0,512949	A1	5.3. Alimentos
	lechuga	0,53807	A2	5.3. Alimentos
	lenteja	0,470055	B1	5.3. Alimentos
	paella	0,532476	A1	5.5. Platos
	pasta	0,584875	A2	5.3. Alimentos
	refresco	0,526194	B2	5.2. Bebida
	tomate	0,615376	A2	5.3. Alimentos
	verdura	0,461414	A1	5.3. Alimentos
Nivel 5	bocadillo	0,284398	A1	5.3. Alimentos
	espagueti	0,2944		
	Fanta	0,297434		
	filete	0,317963	A2	5.3. Alimentos
	hamburguesa	0,327463	A1	5.3. Alimentos
	manzana	0,444405	A2	5.3. Alimentos
	naranja	0,401905	A2	5.3. Alimentos
	pizza	0,357435		
	plátano	0,282128	A2	5.3. Alimentos
	pollo	0,358437	A2	5.3. Alimentos
	potaje	0,326473		

	puchero	0,314485		
	sopa	0,369157	A1	5.5. Platos
	té	0,376824	A1	5.2. Bebida
	tortilla	0,432196	A1	5.5. Platos
	tortilla de patatas	0,288065	En el nivel A1: aparece tortilla	
Nivel 6	aceite	0,196677	B1	5.3. Alimentos
	atún	0,206803	A2	5.3. Alimentos
	azúcar	0,140417	A2	5.3. Alimentos
	batido	0,274646		
	calabaza	0,172616		
	cebolla	0,234217	A2	5.3. Alimentos
	cerdo	0,18687	A2	5.3. Alimentos
	cereal	0,202538	A2	5.3. Alimentos
	chorizo	0,188728	B1	5.3. Alimentos
	cocido	0,180492		
	fresa	0,228662	En el nivel A2 aparece helado de fresa	
	galleta	0,256655	A2	5.3. Alimentos
	garbanzo	0,148535	B1	5.3. Alimentos
	gazpacho	0,180183	Inventario de referentes culturales. Apartado 1.4.1. Platos típicos de las demarcaciones territoriales. <i>Gazpacho</i> (España)	
	güisqui	0,230566		
	jamón	0,256852	A2	5.3. Alimentos
	legumbre	0,214103	B2	5.3. Alimentos
	limón	0,24784	En el nivel C1 aparece corteza de limón	
	macarrón	0,276936		
	maíz	0,165424	En el nivel B2 aparece harina de maíz	
	mantequilla	0,153964	A2	5.3. Alimentos
	melón	0,207278		
	pera	0,196799		
	pimiento	0,204132		
	queso	0,228864	A2	5.3. Alimentos
	ron	0,204901	A2	5.2. Bebida
	sal	0,183262	A2	5.3. Alimentos
	salchicha	0,173112		
	sandía	0,162259		
	solomillo	0,140622	B1	5.3. Alimentos
	tarta	0,229348	A2	5.3. Alimentos
	ternera	0,248955	A2	5.3. Alimentos
	yogur	0,140062	A2	5.3. Alimentos
	zanahoria	0,263661	A2	5.3. Alimentos

Tabla 6.42. Nivel de referencia del léxico nuclear según el PCIC

A simple vista, se observa la relación del léxico compatible con el nivel de usuario básico (A1 y A2) según el MCER. Tan solo se destacan las siguientes unidades léxicas en niveles superiores:

- Nivel B1: *lenteja, aceite, chorizo, garbanzo y solomillo*.
- Nivel B2: *refresco y legumbre*.

El resto de palabras que no se encuentran marcadas en ningún color no aparecen clasificadas en el PCIC.

- Nombres de marcas: *Coca-Cola* y *Fanta*. Cabe destacar que en el apartado dedicado a la red de transportes aparecen los nombres propios *IBERIA* y *RENFE* (Red Nacional de Ferrocarriles Españoles). En cambio, en la alimentación han obviado los nombres de marcas con independencia de la lexicalización de las mismas.
- Vocablos concretos: *espagnetti, pizza, potaje, puchero, batido, calabaza, cocido, güisqui, macarrón, pera, pimienta, salchicha, sandía*.
- Unidades léxicas que aparecen de forma parcial en el inventario: *tortilla* se localiza en el nivel A1 pero no aparece *tortilla de patatas*; *fresa* se ha marcado en naranja porque en el nivel A2 aparece la entrada *helado de fresa*; *gazpacho* aparece como un plato típico en el inventario de referentes culturales sin indicar el nivel según el MCER; *limón* solo aparece con la locución corteza de limón en el nivel C1; *maíz* se incluye en el nivel B2 como *harina de maíz*.

Del cotejo de los inventarios del PCIC y del léxico compatible surgen los siguientes interrogantes: ¿Realmente *aceite* o *refresco* son vocablos que corresponden a un usuario competente? ¿Deberían enseñarse en niveles inferiores? ¿Cómo reconoce un docente el nivel correspondiente a una palabra que no aparece en el PCIC? ¿Cuál es la relación más frecuente entre las palabras de un mismo ámbito y nivel?

6.6.5. Propuesta didáctica

Las preguntas abiertas precedentes sirven de antesala a una propuesta didáctica que aúna necesidades detectadas en la práctica docente con los fundamentos teóricos y empíricos de este estudio. El objetivo es crear una herramienta que permita a los profesionales de la enseñanza obtener información

relevante sobre las unidades léxicas para que puedan hacer una programación didáctica acorde con el nivel de competencia lingüística que desean enseñar. Los aspectos preliminares que se han de tener en cuenta son los siguientes:

- Selección de los centros de interés. Se propone seleccionar las 20 nociones específicas del PCIC ya que son las que guían la elaboración de materiales y posterior evaluación para los Diplomas de Español como Lengua Extranjera del Instituto Cervantes.
- Asignación del nivel correspondiente al MCER según el PCIC. Resulta muy práctico para un profesional de la enseñanza de idiomas conocer a qué nivel del MCER corresponde cada vocablo. Un hablante nativo puede considerar que una palabra es básica porque se emplee con frecuencia en su comunidad de hablante y un docente no hispanohablante puede considerar una palabra más o menos fundamental por su influencia con la lengua materna o por el manual seleccionado y ninguna de ellas pertenecer a un nivel básico.
- Revisión de la clasificación de algunos vocablos según el PCIC. Este paso se realizaría con dos objetivos: completar el nivel según el MCER de las palabras que no aparecen en el PCIC y reasignar otro nivel a algunas lexías como por ejemplo aceite, que se encuentra como palabra compatible en los tres idiomas pero aparece en el nivel B1 en el PCIC. Este vocablo no aparece en los listados en inglés y francés pero en la cultura gastronómica española es esencial.
- Determinación de la noción o nociones de la que forma parte. Esta información es imprescindible para conocer en qué ámbito o tema se suele enmarcar esa unidad léxica.
- Cálculo de la disponibilidad léxica. Este índice muestra si la palabra es accesible a los informantes mediante una fórmula que incluye la posición de la palabra en la lista y el número de informantes que la ha escrito.
- Cálculo del índice de compatibilidad. Este indicador muestra la distancia de la palabra con respecto al prototipo.
- Nivel de compatibilidad. Destacan las palabras básicas divididas en 6 niveles que suelen corresponder los niveles A1 y A2.

- Relaciones más frecuentes entre las palabras. Muestra de los *clusters* y relaciones que los informantes nativos realizan. Por ejemplo: *carne* y *pescado* cuenta con 47 menciones en uno u otro orden.
- Porcentaje de conectividad. Informa sobre el papel que ocupa un vocablo con respecto a la red de conexiones de ese centro de interés. Cuanto mayor sea el índice mejor será la capacidad de relación de esa palabra. Se calcula analizando el total de relaciones de un centro entre las relaciones de esa palabra y multiplicándolo por 100.
- Adecuación a los objetivos de aprendizaje. Si algunos vocablos pertenecen a un vocabulario técnico o especializado se debería señalar además de por el nivel según el MCER para las lenguas con una referencia ya que puede tratarse de una palabra básica en un ámbito determinado pero al mismo tiempo técnica. También se debe hacer constancia de las palabras que provengan de una región o lugar concreto. Por ejemplo, si se marca *gazpacho* como un plato típico de España seguramente un docente mexicano no lo enseñará a sus estudiantes en un nivel básico al igual que en España no se incluirá *pibil* en los manuales.

Agua	
Centro de interés	5. Alimentación
Nivel según el MCER	A1
Noción específica	5.2. Bebida
Nivel de compatibilidad	1
Índice de compatibilidad	0,955196
Rango	1 de 530
Índice de disponibilidad léxica	0,44275
Rango	1 de 530
Porcentaje de aparición	60,234
Porcentaje de conectividad con respecto al centro	8,91 %
Rango	1
Número total de conexiones	181
Número de conexiones desde el término	102
Conexiones más frecuentes desde el término	leche (11), zumo (11), vino (11), refresco (10), cerveza (9)
Número de conexiones hacia el término	79
Conexiones más frecuentes hacia el término	leche (8), pan (7), zumo (6), Fanta (5), vino (4)
Enlace a la entrada en el DRAE	http://lema.rae.es/drae/?val=agua

Tabla 6.43. Ejemplo de la búsqueda del vocablo agua en la base de datos

Por último, cabe destacar que se propone la selección de las 20 nociones específicas del PCIC para obtener un diccionario de léxico disponible general. En

cambio, si el objetivo es elaborar un listado léxico con fines específicos sería conveniente tomar como estímulo los subcampos relacionados. Por ejemplo, si el propósito es enseñar español administrativo o económico sería conveniente obtener el léxico a partir de los subcampos *Financias y bolsa, renta, comercio, entidades y empresas*. En este caso, es imprescindible seleccionar una muestra conocedora del tema y con una competencia léxica suficiente para obtener un léxico nuclear acorde a las necesidades de los estudiantes.

6.7. Conclusiones parciales

A lo largo de este capítulo, se ha realizado un recorrido desde la teoría sobre la organización de las palabras en la mente, pasando por análisis sobre compatibilidad léxica y relaciones entre vocablos, hasta llegar a una propuesta práctica de enseñanza basada en el léxico disponible. De esta forma, la aplicación pedagógica se sirve de unos fundamentos teóricos sólidos y unos análisis que lo corroboran. A colación de ello, cabe citar al pintor, escultor e inventor italiano Leonardo Da Vinci (1452-1519): «Los que se enamoran de la práctica sin la teoría son como los pilotos sin timón ni brújula, que nunca podrán saber a dónde van».

Al inicio del trayecto, el aguja de la brújula indica cuáles son las teorías que ayudan a explicar mejor los procesos cognitivos que tienen lugar en la mente del hablante cuando realiza la prueba de disponibilidad léxica: la teoría de los prototipos y la teoría de las redes semánticas. Siguiendo ese orden, se proporciona el vocabulario prototípico de cada centro de interés basado en la aplicación del modelo de conjunto difusos a la compatibilidad léxica que realizan Ávila Muñoz y Sánchez Sáez (2010, 2011, 2014). Se ha comprobado el alto grado de similitud entre el léxico nuclear en español, inglés y francés que ratifica que las palabras se sitúan en torno a un núcleo o prototipo. Se convierte así en una herramienta clave para aunar criterios en los estudios de disponibilidad léxica en la selección del vocabulario básico que se corresponde con los niveles A1 y A2 según la clasificación del Marco Común Europeo de Referencia para las Lenguas.

En el tercer apartado, se ha realizado un alto en el camino para conocer cuáles son los procesos de reentrada al núcleo que se producen en las secuencias léxicas

aportadas por cada informante y si se ven influenciadas por el tipo de centro. Se ha comprobado que desde los inicios de la disponibilidad léxica, los autores franceses Gougenheim *et al.* (1964: 189-190) destacaron el diferente comportamiento de los centros de interés según el número de palabras y vocablos aportados por los informantes. *A priori*, distinguieron entre niveles de homogeneidad y pusieron de manifiesto la mayor idoneidad de unos centros con respecto a otros debido a los niveles de disponibilidad que presentaban.

Esa distinción está presente en la mayoría de los resultados comparativos generales de las investigaciones donde se organizan los centros según el número de palabras y vocablos aportados y el índice de cohesión (Carcedo González, 2000c; Samper Hernández, 2002; Ávila Muñoz y Villena Ponsoda, 2010; Bellón Fernández, 2011, entre otros). Algunos estudios como los de Hernández Muñoz (2006), Tomé Cornejo (2011) y Sánchez-Saus Laserna (2011) presentan una clasificación atendiendo a la estructura interna de los centros.

Tras analizar las taxonomías de los estudios anteriores y los diferentes resultados generales obtenidos en el Capítulo 4, se ha propuesto una clasificación de los centros de interés según su prototipicidad. El objetivo es conocer el comportamiento de cada uno a través del compendio de tres índices: el índice de cohesión, que mide la relación del promedio de palabras de un informante en un centro entre los vocablos de ese centro; el índice de centralidad que muestra la distribución de los vocablos nucleares con respecto a los vocablos totales de un centro; y el índice de conectividad que da cuenta de la mayor o menor solidez de las relaciones de los vocablos de un centro. Este último se obtiene restando uno a la cantidad resultante de dividir el número de conexiones diferentes entre el total de palabras.

El resultado es un índice estructural que da cuenta de la homogeneidad del centro, de la organización de las palabras nucleares y de la fuerza de las conexiones entre los vocablos. Esta expresión numérica servirá para organizar los centros en una escala cuyos extremos están formados por el centro más prototípico, *Partes del cuerpo humano*, y el centro más relacional o lo que es lo mismo, menos prototípico, *Ordenadores e internet*. El resto de los centros se posiciona en ese continuo que ayuda a

definir su estructura. Así, *La ropa y Medios de transporte* se acercan al grupo de centros prototípicos, *La Escuela: muebles y materiales* se ubica en un nivel intermedio y el resto de centros, *Comidas y bebidas*, *La ciudad*, *Juegos y distracciones* y *Profesiones y oficios* muestran resultados semejantes al centro más relacional.

Esa clasificación es muy similar en los tres idiomas del estudio e incluso podría decirse que será muy parecida a otros estudios. En cambio, esta propuesta puede verse alterada por factores psicolingüísticos y sociolingüísticos por lo que dependerá en gran parte de las respuestas de cada muestra. Este esbozo de la taxonomía de los centros viene a completar las clasificaciones teóricas con resultados prácticos ya que este estudio no se ha limitado a incluir los vocablos que mantienen una relación directa con el centro de interés sino que se han incorporado todas las producciones de los individuos con independencia del grado de relación con el prototipo. A continuación, se presentan los procesos de reentrada al núcleo según el tipo de centro de interés mostrando un diferente comportamiento en los centros más y menos prototípicos.

- Centros prototípicos

En este tipo de centros se han detectado tres tipos de modelos que forman parte de un continuo. En un extremo se situaría el patrón de reentrada leve que muestra el comportamiento de las secuencias léxicas de los informantes con un caudal léxico reducido. En líneas generales, se trata de palabras muy compatibles que apenas se alejan del núcleo. En el extremo derecho se sitúa el modelo de reentrada pronunciado que muestra desvíos abruptos entre una unidad léxica y otra. Este tipo de modelos corresponden a informantes con un caudal léxico rico y variado que son capaces de escribir una palabra común y, seguidamente, otra específica. En general, esas vocablos sirven para definir con más detalles los subcampos dentro de cada centro de interés aunque también pueden ser fruto de asociaciones individuales.

En el centro de esa escala se encuentra el modelo de reentrada moderado en el que se suceden desviaciones del núcleo leves, moderadas y pronunciadas. La forma de los desplazamientos con respecto al núcleo va moldeándose conforme el informante posee mayor o menor riqueza hasta acercarse a uno de los dos extremos. Las respuestas suelen incluir una serie de palabras nucleares que dan lugar a

reentradas suaves y vocablos más específicos que ayudan a definir mejor la estructura o prototipo del centro de interés y dibujan una reentrada más abrupta.

Cabe mencionar que en lengua española y en lengua extranjera los modelos son muy similares ya que se ven influenciados por la riqueza del vocabulario del informante. En el caso de las respuestas en inglés y en francés, se observa que en el proceso de reentrada leve la organización léxica no muestra una estructura clara en la mente del informante, sino que se limita a aportar las palabras que conoce, razón por la cual algunas no guardan relación entre sí. En cambio, cuanto mayor es el léxico disponible de los hablantes las respuestas en lengua extranjera y lengua materna se neutralizan, llegando incluso a reproducir partes y estructuras anteriormente citadas.

Este fenómeno se ha constatado en el informante 38 que cuenta con el mayor número de respuestas en francés en el primer centro de interés, concretamente 49, mientras que en español desciende a 45. Entre los motivos de tales resultados, destaca la propia metodología de la prueba realizada, ya que primero se realiza en español y luego en lengua extranjera. El sujeto bilingüe es capaz de reflexionar sobre un centro de interés en español y, posteriormente, trasladar esa misma estructura a la lengua extranjera reproduciendo algunas secuencias léxicas.

- Centros relacionales

En los centros relacionales se observan dos modelos que marcan la escala entre los informantes con un repertorio lingüístico reducido y aquellos con un amplio caudal léxico. En el primer caso, los informantes que han aportado pocas palabras dibujan un modelo de reentrada moderado cuyas palabras suelen pertenecer al listado de palabras nucleares definida mediante el modelo de conjuntos difusos. A diferencia de los centros inclusivos, en este tipo de centros el límite del núcleo está más disperso, lo que lleva a los informantes con poco caudal léxico a desviarse ligeramente más del núcleo.

En la parte opuesta, se sitúan los informantes que han aportado muchas palabras. El tipo de relación que se realiza es de tipo horizontal referida a algún subcampo léxico. Por ejemplo, en el caso del centro de interés *La escuela: muebles y materiales* los subcampos léxicos podrían ser: material escolar, muebles, asignaturas, etc. Las relaciones de meronimia no son frecuentes y el repertorio lingüístico es más

amplio lo que provoca que los procesos de reentrada hacia términos específicos no sean tan frecuentes y que los informantes se desplacen del núcleo con asociaciones de diversa índole.

- Clasificación del centro de interés *Ordenadores e internet*

Se ha dedicado un apartado a describir los procesos de reentrada del centro de interés denominado *Ordenadores e internet* porque no se encuentra descrito en otros estudios (Hernández Muñoz, 2006; Sánchez-Sauz Laserna, 2011) y que se define como el centro más relacional según el índice estructural. Tras analizar las respuestas de los informantes se han detectado dos tipos de procesos de reentrada: moderado y pronunciado con desviación lineal. Esos modelos corresponden a los descritos en los centros de interés radiales.

La organización de las palabras de los informantes en torno al núcleo no es más que otra prueba que muestra que el centro no tiene una estructura interna ni bien definida ni natural. El informante se enfrenta a un centro heterogéneo (como muestra el índice de cohesión) donde apenas conoce los límites de representatividad. Por esa razón, es el centro que más nombres propios aporta ya que en lugar de relacionar internet con aspectos técnicos y de funcionamiento lo asocian con diferentes páginas web que visitan.

En este tipo de centros, la relación se realiza sobre todo a nivel horizontal y los subcampos se encuentran algo más dispersos en la secuencia léxica. Es frecuente encontrar relaciones entre pares de palabras que el informante relaciona con otras con las que guarda cierta relación directa, indirecta e incluso personal. El hecho que el centro no esté tan bien acotado conlleva que el informante que aporta una gran cantidad de palabras se desvíe progresivamente del núcleo y cuando vuelve lo haga con menos intensidad que en el caso de los centros inclusivos.

Los procesos de reentrada descritos confirman que el tipo de centro de interés condicionará las respuestas de los informantes tanto en el número de palabras aportadas como en la compatibilidad de las mismas. Ello explica que la selección del léxico se haga mediante la teoría de conjuntos difusos y no simplemente seleccionando los vocablos superiores a algún índice concreto o entre unas posiciones determinadas. En esos casos, quizás algunos centros cuenten con un

excedente de palabras que provoque que algunas no sean tan compatibles y en otros en que falten palabras consideradas nucleares.

Una vez seleccionado el léxico nuclear y analizados los procesos de reentrada al núcleo, se han analizado las conexiones más frecuentes entre las palabras. Los análisis muestran las relaciones semánticas generadas a partir del programa DispoGrafo en los tres idiomas del estudio para tratar de responder a varias hipótesis iniciales. Por un lado, se ha pretendido reafirmar la existencia de centros de diferente naturaleza según las descripciones realizadas por Hernández Muñoz (2006: 150-156) y Sánchez Saus-Laserna (2011: 196-209) y su incidencia en los resultados de cada centro. Para ello, se analizan datos generales como el número de palabras y vocablos, la densidad prototípica, la densidad de conectividad y la cantidad y complejidad de los *clusters* ante estímulos diferentes: *Partes del cuerpo humano*, *La escuela: muebles y materiales* y *Ordenadores e internet*.

Se corrobora la hipótesis que la ontogenia de los centros de interés marcan los resultados. Así, los centros más compactos dan lugar a relaciones semánticas más densas, como es el caso del centro *Partes del cuerpo humano* y aquellos más heterogéneos establecen relaciones semánticas más dispersas, como el centro *Ordenadores e internet*. La influencia del centro en el número de palabras y vocablos conlleva una alteración en el número de relaciones y en la fuerza de las mismas. Según Borrego Nieto (2004: 67):

Existen relaciones heterogéneas entre el enunciado del campo y el léxico resultante: en *Partes del cuerpo* o en *Animales* se obtienen listas de partes del cuerpo o de animales, es decir, los enunciados son hiperónimos de las palabras que constituyen las respuestas. En cambio, en *El campo* y *La ciudad*, se obtienen los vocablos correspondientes a las realidades que evocan, y no hipónimos de los enunciados.

Por otro lado, se realizan comparaciones léxicas intermuestrales entre los tres idiomas del estudio (inglés, francés y español) para observar las principales semejanzas y diferencias en la organización del léxico disponible. En este sentido se concluye, igual que hicieron Ferreira (2006) y Ferreira y Echeverría (2010-2011), que los núcleos semánticos están mejor definidos en la lengua materna de los informantes que en lengua extranjera. Asimismo, se constata que las relaciones semánticas son más densas en español, factor que se ve influenciado por el mayor número de vocablos.

Una vez seleccionado el léxico nuclear y analizadas las relaciones semánticas según el tipo de centro de interés, se ha estudiado la importancia de los aspectos socioculturales en los resultados del léxico disponible. Se trata de demostrar que las unidades léxicas aportadas dependen de la procedencia de los informantes, su edad y nivel de instrucción, entre otros aspectos. Por esta razón, es imprescindible conocer las características de la muestra de cada léxico disponible y saber si se adecuan a los objetivos de enseñanza.

El siguiente paso para realizar una propuesta didáctica ha consistido en una revisión de los centros de interés. Con anterioridad, otros autores como Bartol Hernández (2010), Sánchez-Saus Laserna (2011), Paredes García (2014) han propuesto renovar los dieciséis centros tradicionales por otros más adaptados a la situación actual. Es bien sabido que existen centros poco productivos, como *Trabajos del campo y del Jardín e Iluminación, calefacción y medios de airear un recinto*, y que existen lagunas en los diccionarios del léxico disponible que dejan de lado aspectos básicos como la dimensión perceptiva y anímica del individuo.

Dada la importancia de la publicación del Marco Común Europeo de Referencia para las Lenguas por el Consejo de Europa en 2001³¹⁴, se ha comparado la propuesta de nociones específicas presente en el Plan Curricular del Instituto Cervantes con los centros de interés. Se ha constatado que los centros de interés solo cubren algunas parcelas de las nociones imprescindibles en el aprendizaje de una lengua. Además, los centros de interés que autores como Paredes García (2014) proponen sustituir se encuentran dispersos en la clasificación del PCIC.

A continuación, se han comparado los listados del léxico nuclear o más compatible en los tres idiomas estudiados. A pesar de que el número de lexías nucleares se ve influenciado por el número de informantes, se ha podido comprobar la gran similitud entre el léxico nuclear de los listados en lengua materna y lengua extranjera. En inglés tan solo se incluyen 5 palabras que no están presentes en el listado español y en francés se mencionan 10 diferentes. No obstante, la mayoría aparece en el diccionario de léxico disponible en español pero en posiciones inferiores y con un índice de compatibilidad más bajo. Estos resultados constatan la

³¹⁴ La publicación en español se realiza un año más tarde (Consejo de Europa, 2002).

presencia de un prototipo en cada centro de interés con independencia del idioma de estudio. También se corrobora la incidencia de los factores socioculturales en los listados del léxico, con palabras más propias de una cultura gastronómica que de otra: *choucroute* en francés, *bean* en inglés o *paella* en español.

A partir de las respuestas de los informantes en el centro de interés *Comidas y bebidas*, se ha localizado el nivel según el MCER al que corresponde cada unidad léxica que conforma el listado del léxico compatible. Como era de esperar, la mayoría de vocablos pertenecen a los niveles de usuario básico (A1 y A2) confirmando la hipótesis que el léxico nuclear fija los límites del vocabulario básico de una lengua. En cambio, existen ciertas palabras como *aceite* que el PCIC propone en un nivel B1 y cuenta con una compatibilidad equiparable a un A2. El PCIC se basa en clasificaciones realizadas en otros idiomas y, por tanto, herederas de otras culturas culinarias donde el aceite quizás no juega el mismo papel que en la gastronomía española.

Aún así, no todas las palabras han podido ser catalogadas según el nivel del MCERL porque no aparecen en el PCIC. Para salvaguardar ese inconveniente, se propone la creación de una base de datos con fines didácticos que sirva a los profesionales de la educación a seleccionar adecuadamente el léxico según el nivel de los estudiantes y los objetivos del aprendizaje de la lengua. De esta forma, el estudiante se convierte en el centro del proceso de enseñanza-aprendizaje y se consigue un aprendizaje significativo adaptado a sus necesidades. La propuesta informática se concreta en una ficha que ilustra la información que debería contener de cada unidad léxica: nivel según el MCER, nombre del centro de interés, noción específica del PCIC, índice de disponibilidad léxica, índice de compatibilidad léxica, relaciones más frecuentes entre palabras y un enlace a la información del Diccionario de la Real Academia Española.

CAPÍTULO 7. CONCLUSIONES

7.1. Conclusiones

7.1.1. Ámbito de la sociolingüística

7.1.1.1. Análisis general

7.1.1.2. Análisis sociolingüístico

7.1.2. Ámbito de la psicolingüística

7.1.3. Ámbito de la lingüística aplicada a la enseñanza de idiomas

7.1.4. Propuestas y líneas de investigación futuras

7.2. Conclusions

7.2.1. Cadre de la sociolinguistique

7.2.1.1. Analyse générale

7.2.1.2. Analyse sociolinguistique

7.2.2. Cadre de la psycholinguistique

7.2.3. Domaine de la linguistique appliquée à l'enseignement des langues

7.2.4. Propositions et lignes de recherches futures

7.1. Conclusiones

Al igual que los objetivos e hipótesis, las conclusiones se presentan según el ámbito de aplicación. En ciertas ocasiones resulta difícil discernir la línea que separa cada disciplina ya que mantienen una relación simbiótica en la que una complementa a la otra. Así, las diferentes respuestas del vocabulario en español y en lengua extranjera nos permite extraer conclusiones que tienen una aplicación directa en los programas educativos. Además, si el interés está orientado a conocer la influencia que ejercen las diferentes variables sociales sobre el léxico disponible estaría dentro del campo de la sociolingüística. Por último, se han incluido algunas propuestas y líneas de investigación futuras generadas a partir de los resultados de este trabajo.

7.1.1. Ámbito de la sociolingüística

En el ámbito de la sociolingüística se ha pretendido conocer cuáles son los factores que inciden en el vocabulario de los informantes haciendo una distinción inicial entre vocabulario activo y pasivo. Por un lado, el vocabulario activo se refiere a aquel que el informante es capaz de utilizar en contexto de forma adecuada y se obtiene por medio de la prueba de léxico disponible. Por otro lado, el vocabulario pasivo se entiende como aquel que el informante es capaz de reconocer y se ha asociado a la prueba de identificación de tecnicismos.

En primer lugar, se presentan los resultados generales de una indagación exploratoria sobre el estudio del léxico activo desde dos perspectivas complementarias: intramuestral e intermuestral. A continuación, esos datos sirven de punto de partida de las conclusiones de un examen estadístico donde se evalúa la incidencia de variables sociológicas y relacionadas con el aprendizaje de una lengua sobre el léxico activo y pasivo en español y en lengua extranjera.

7.1.1.1. Análisis general

Las indagaciones del léxico disponible en español, inglés y francés de los estudiantes del Máster en Profesorado dan cuenta de una considerable diferencia

entre el caudal léxico en lengua materna y en lengua extranjera. En español, los informantes han escrito un promedio de 191,91 palabras, en inglés 131,79 y en francés 147,71. Las comparaciones con estudios previos corroboran, además de la hipótesis del mayor caudal léxico en los casos de lengua materna, la influencia del nivel de estudios en el léxico disponible y en la correcta representación ortográfica.

En español como lengua extranjera, los estudiantes finlandeses (Carcedo González, 2000c) aportaron una media de 13,68 palabras por centro de interés; los extranjeros que realizaron una estancia en Salamanca (Samper Hernández, 2002) superaron esa media con un promedio de 15,63 palabras. En cambio, los españoles del estudio actual mejoraron ese promedio con 17,24 unidades léxicas en inglés y francés. Por tanto, el caudal léxico alcanzado en lengua extranjera por los estudiantes españoles es considerablemente más alto que el de los estudiantes extranjeros de español. Entre los factores que favorecen esa mayor capacidad léxica se puede destacar el nivel de instrucción, puesto que todos los informantes han debido acreditar un mínimo de un B1 y estar en posesión de un título universitario mientras que en los estudios con informantes extranjeros el nivel de competencia lingüística podía ser inferior.

Resultados similares se obtienen en las comparaciones con estudios que analizan el vocabulario disponible en inglés. En la investigación llevada a cabo por Ferreira (2006), se analiza el léxico actualizado por estudiantes chilenos de Pedagogía en Inglés y Traducción en Lengua Extranjera y de escolares ingleses. El promedio de palabras aportadas por los informantes nativos es de 27,28, por los estudiantes chilenos 18,69 y por los estudiantes de postgrado de este estudio es de 23,19³¹⁵. Del mismo modo, los escolares escoceses de la investigación de Dimitrijević (1969) aportaron una media de 20,04 palabras en inglés mientras que la media de los estudiantes de este estudio en inglés es de 17,69 palabras. Por tanto, el factor que guarda una mayor relación con el caudal léxico es la competencia lingüística en el idioma de la prueba ya que los estudiantes nativos actualizan una mayor cantidad de palabras.

³¹⁵ La comparación con este estudio, al igual que con los anteriores, se ha realizado contabilizando la media de palabras en los centros de interés comunes.

La incidencia directa del nivel educativo en el promedio de palabras se analiza a través del cotejo de los resultados de estudios llevados a cabo en Málaga con estudiantes preuniversitarios (Ávila Muñoz, 2006), con población adulta preestratificada (Ávila Muñoz y Villena Ponsoda, 2010) y con los estudiantes de postgrado. La media de palabras por centro de interés calculada sobre los ocho centros comunes es de 19,69 en los estudiantes preuniversitarios, 20,50 en la población adulta estratificada y 21,59 en los estudiantes de postgrado. Estos resultados casan con los obtenidos previamente del cotejo del léxico disponible de los universitarios chilenos donde se mostraba que su caudal léxico en inglés es menor que el correspondiente a los estudiantes de postgrado españoles.

Con respecto a la corrección ortográfica, se comprueba que los estudiantes de postgrado han cometido menos errores que los estudiantes preuniversitarios, lo que parece lógico si se considera su paso por la universidad. Mientras que los preuniversitarios madrileños (Paredes García, 1999) han cometido una media de 2,26 faltas de ortografía por informante y centro de interés y los preuniversitarios aragoneses 1,83 (Saura Rami, 2008), los estudiantes de postgrado tan solo 0,28 en español y 0,73 en lengua extranjera. En cuanto al porcentaje de errores registrados en las encuestas de los universitarios malagueños (Ávila Muñoz, 2007), se observa que es considerablemente superior al registrado en los estudiantes de postgrado. Por ejemplo, tan solo un 6 % de los informantes universitarios no han cometido ningún error ortográfico frente al 21,05 % de los estudiantes de postgrado en español. Como se postulaba en las hipótesis iniciales, en lengua extranjera el porcentaje de corrección ortográfica de los postgraduados es menor que en lengua materna y desciende al 3,51 %.

El análisis de la situación ortográfica ha permitido detectar cuáles son los errores ortográficos más usuales en los estudiantes con objeto de extraer conclusiones didáctico-críticas y consolidar resultados de estudios previos (Paredes García, 1999; Ávila Muñoz, 2007 y Saura Rami, 2008). Los resultados revelan que del total de faltas de ortografía, el 66,16 % de las disortografías corresponden a una inadecuada acentuación de las palabras, ya sea porque se ha acentuado cuando no debiere o viceversa. A continuación, se sitúa el uso incorrecto de los grafemas *b* y *v* ya

que la poligrafía del fonema /b/ provoca confusiones, al igual que la diferente escritura del fonema /x/.

En cuanto a las unidades léxicas más disponibles, se han analizado los veinte primeros vocablos de varios centros de interés comunes en diversos estudios (Gougenheim *et al.* 1964; Dimitrijević, 1969; Carcedo González, 2000c; Samper Hernández, 2002; Ferreira, 2006; Ávila Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010). En general, se ha constatado una mayor similitud entre los listados de estudios que evalúan el léxico de una lengua extranjera con respecto a investigaciones donde la lengua analizada es la materna de los informantes. De esta forma, entre los vocablos más disponibles en lengua materna los informantes han mencionado cuatro relacionados con la ropa interior mientras que en lengua extranjera solo han aportado una o dos. Ese tipo de prendas se usan diariamente pero no aparecen con tanta frecuencia en los manuales de inglés, francés o español como lengua extranjera.

Asimismo, el espacio temporal de recogida de datos muestra la influencia de la diacronía de la lengua condicionando el vocabulario más disponible. De esta forma, en el estudio pionero de Gougenheim *et al.* (1964) aparecen vocablos como *combinaison* que probablemente hoy en día no serían tan disponibles por dejar de ser una prenda tan habitual. Por tanto, es posible observar la relación entre el uso y el contexto de aprendizaje en el léxico disponible que será un factor determinante en la selección del léxico con fines educativos.

7.1.1.2. Análisis sociolingüístico

La variable utilizada para medir la capacidad léxica ha sido el índice corregido de descentralización léxica (ICoDD). El objetivo de la descentralización léxica es conocer la especificidad de las palabras aportadas por un informante con respecto a las palabras escritas por el resto de sujetos. Para evitar que este índice otorgue valores superiores a informantes que han mencionado palabras fruto de asociaciones individuales, los vocablos se dividen en cuatro tipos atendiendo a la relación con el núcleo o prototipo del centro de interés que corresponde a la clasificación realizada en el índice corregido de amplitud léxica.

El índice corregido de descentralización léxica se correlaciona significativamente con las variables que hacen referencia a la suma de todas las palabras aportadas. En este caso se distingue el índice de amplitud léxica que otorga el valor de 1 a todas los vocablos y el índice corregido de amplitud léxica que toma en cuenta la relación de la palabra con el centro de interés y establece cuatro tipos de categorías de palabras (N1, N5, N25 y N0). Por tanto, el ICoDD sirve para medir la capacidad léxica de los informantes según la relevancia de las palabras en el centro de interés y su carácter restringido. Se presupone que aquellos informantes que son capaces de escribir palabras pertenecientes a un ámbito más especializado, como *calesa* en el centro *Medios de transporte*, serán capaces de escribir palabras compartidas por la mayoría de informantes, como puede ser *coche*.

La capacidad de identificación del léxico especializado se obtiene a partir de una prueba que consiste en relacionar palabras técnicas o especializadas con uno de los nueve centros de interés que se incorporan en este estudio. La variable resultante tiene en cuenta el valor azaroso e intrínseco en todas las pruebas de elección múltiple y otorga un valor diferenciado a las respuestas correctas e incorrectas. A continuación, se resta el cómputo de las respuestas incorrectas ponderado a las respuestas correctas³¹⁶.

En el ICoDD en lengua extranjera la variable con mayor relevancia en el modelo de regresión lineal y causal es la especialidad del máster. Las especialidades con una influencia positiva en el caudal léxico son Lengua Extranjera Inglés, Francés, Lengua y Literatura, Latín y Griego. Las especialidades que se mantienen en la media son Geografía e Historia; Dibujo, Imagen y Artes Plásticas; Música, Formación y Orientación Laboral, Matemáticas, Orientación Educativa, Hostelería y Turismo; Tecnología, Informática y Procesos Industriales.

Por orden de importancia, le sigue el caudal léxico en español que guarda una relación positiva con el léxico disponible en inglés y francés. Por tanto, se constata que los resultados en lengua materna inciden en el léxico en lengua extranjera de forma significativa y viceversa. De forma secundaria, se podrían destacar la influencia

³¹⁶*Vid.* apartado 2.5.4 para más información sobre el cálculo de la variable.

indirecta sobre el ICoDD en lengua extranjera de aquellas variables que influyen de forma directa en el ICoDD en lengua española y viceversa.

El grado de motivación del futuro profesorado es un factor clave no solo porque influye directamente en el caudal léxico en lengua extranjera sino también porque se relaciona con otras variables, como puede ser el interés por seguir aprendiendo idiomas, la intención de trasladarse al extranjero, un mayor uso de la lengua, la especialidad del máster, entre otros aspectos. El uso de la lengua extranjera en diferentes contextos es el resultado de un análisis de componentes principales que recoge la información común de las variables relacionadas con la realización de la Erasmus y con la asiduidad del uso de la lengua en el ambiente académico, laboral y familiar.

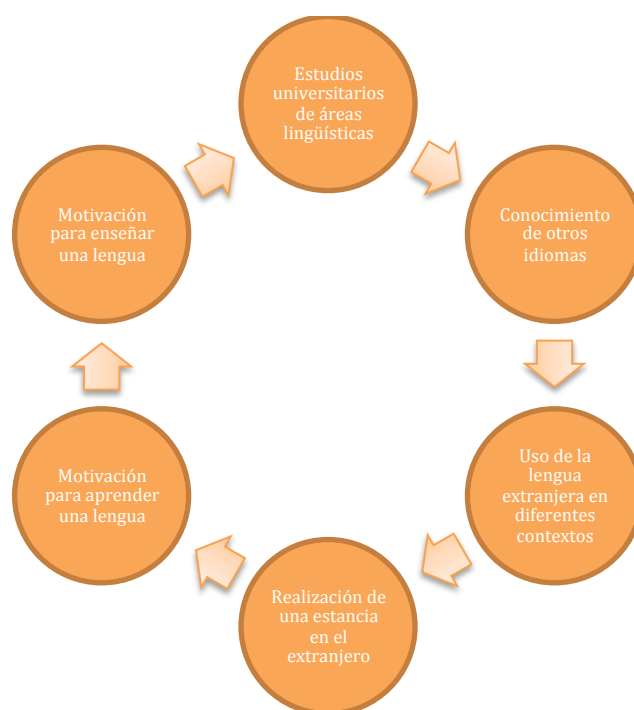


Gráfico 7.1. Resumen de las variables con mayor incidencia en el léxico en lengua extranjera

Los modelos multicausales en lengua española explican con menor precisión la incidencia de variables sobre el caudal léxico. Se entiende que el factor con mayor peso es el nivel de instrucción alcanzado por los informantes que está condicionado de forma secundaria por otras variables. Las dos variables con mayor peso coinciden con las mencionadas en el modelo en lengua extranjera: el caudal léxico en lengua extranjera y la especialidad del máster. De esta forma, se constata que los sujetos que

tienen un mayor léxico disponible en una lengua lo transmiten a la otra lengua gracias a la flexibilidad cognitiva que presentan los hablantes bilingües.

El comportamiento de las especialidades varía con respecto a la lengua extranjera: Matemáticas es la única especialidad que influye de forma negativa; Lengua Extranjera Inglés y Francés, Lengua y Literatura; Latín y Griego y Música tienen una incidencia positiva y el grupo formado por las especialidades de Geografía e Historia, Dibujo, Orientación Educativa, Formación y Orientación Laboral, Hostelería y Turismo, Tecnología, Informática y Procesos Industriales.

La identificación de tecnicismos se presenta como una variable que favorece la capacidad léxica, lo cual explica que el bagaje en el vocabulario pasivo influya en el vocabulario activo. Este vocabulario se adquiere a través de la lectura que tiene una repercusión menor en el modelo, pero que resulta relevante ya que puede marcar la diferencia en el caudal léxico en hablantes postgraduados.

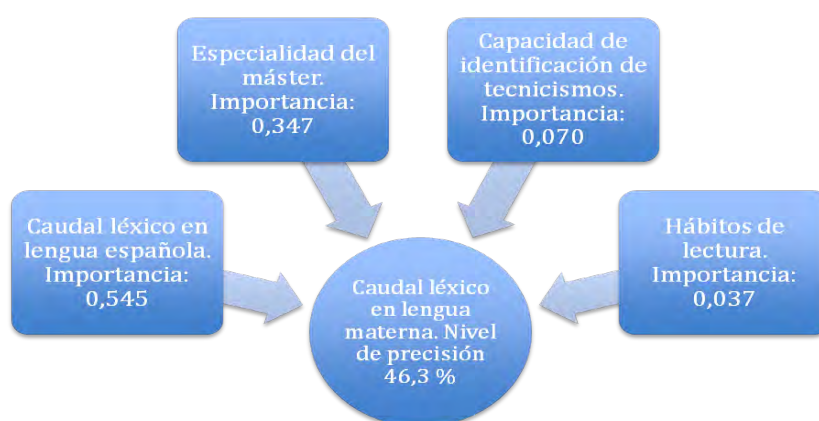


Gráfico 7.2. Resumen del modelo de regresión lineal del caudal léxico en español

Con respecto a las variables que inciden en la identificación del léxico especializado en lengua extranjera y en español coinciden dos: el uso de la lengua en diferentes contextos y el tipo de prueba. Ello conlleva que los informantes con un mayor dominio de las palabras técnicas en lengua española también sean capaces de identificar el léxico especializado en lengua extranjera más fácilmente. Además, el ICIT guarda una relación significativa con el caudal léxico y con la lengua de la prueba. En lengua inglesa los informantes han bajado la media de los resultados, mientras que en francés aumenta ligeramente. Por tanto, la lengua elegida tiene cierta repercusión en los resultados finales de la prueba.

Por último, la clase social influye de forma directa e indirecta en el caudal léxico. Se ha recogido en el modelo causal su influencia a través de la variable del uso de la lengua y del caudal léxico en lengua española. Además, aparece con mayor influencia la especialidad en lengua extranjera que está relacionada con el número de idiomas que habla un informante y el nivel de competencia lingüística y el uso de la lengua extranjera en diferentes contextos, entre otros aspectos.

Los resultados en español muestran un mayor conocimiento del léxico especializado que en lengua extranjera. La especialidad presenta una incidencia diferente. En este caso, aquellas variables con una influencia positiva son Lengua Extranjera Francés y Tecnología, Informática y Procesos Industriales. La única especialidad que hace descender el ICIT es Matemáticas. El resto de especialidades (Lengua Extranjera Inglés, Lengua y Literatura, Latín y Griego; Dibujo, Imagen y Artes Plásticas; Orientación Educativa; Formación y Orientación Laboral; Hostelería y Turismo; Geografía e Historia) se mantiene en torno a la media.

Al igual que en lengua extranjera, el caudal léxico en español tiene una influencia directa. Asimismo, el léxico especializado en lengua extranjera se muestra relevante por lo que el conocimiento de otros idiomas incide no solo en las pruebas en lengua extranjera sino también ayuda a fomentar el dominio lingüístico en la lengua materna.

Las variables referidas a los hábitos de lectura y la exposición a los programas culturales presentan una influencia indirecta a través de la clase social y directa sobre el léxico disponible. Este resultado conlleva una gran repercusión en el ámbito de la didáctica puesto que sugiere, por un lado, que la lectura no ha calado de igual forma en todas las clases sociales y, por otro lado, que ayuda a mejorar la competencia léxica de los estudiantes. La influencia de la clase social en el modelo causal en el ICIT en español es mayor que en el resto de variables dependientes (ICIT en lengua extranjera e ICoDD en español y en lengua extranjera).

A modo de resumen, se presentan las variables con mayor influencia en el léxico activo y pasivo que han contestado a las hipótesis de estudio. Cabe destacar que no se han incluido aquellas que no ayudan a explicar el comportamiento léxico de los informantes, como el sexo. Al igual que en el estudio de Ávila Muñoz y Villena

Ponsoda (2010: 285) no ha resultado una variable relevante ya que las desigualdades en los roles masculinos y femeninos desaparecen a medida que mejora la situación socioeconómica de las personas.

- Grado de instrucción:

Con respecto a las encuestas de disponibilidad léxica con estudiantes preuniversitarios, se produce un aumento considerable en el número de palabras aportadas por los estudiantes de postgrado de este estudio. Ese grado de instrucción merma la importancia de variables presentes en otras investigaciones como el tipo de centro educativo ya que los estudios especializados hace que se difuminen esas diferencias.

- Conocimiento de otras lenguas:

Los datos aportados por informantes que hablan más de un idioma extranjero muestran un mejor caudal léxico en español y, especialmente, en lengua extranjera. Resulta imprescindible realizar un análisis más pormenorizado sobre la competencia lingüística que han alcanzado los hablantes en cada lengua extranjera ya que incide significativamente en los resultados. De esta forma, informantes que conocen una lengua con un nivel B2 cuentan con una capacidad léxica mayor que aquellos que conocen dos lenguas con niveles inicial y medio ya que generalmente ese nivel B2 corresponde a la lengua que han utilizado para realizar la prueba (inglés o francés).

- Estancias en el extranjero:

Las estancias se han dividido según su duración y financiación distinguiendo entre realización de la Erasmus, otros estudios becados en el extranjero y estudios sin beca. Sin duda, el programa Erasmus muestra una mayor relación con el caudal léxico y especializado en lengua extranjera en comparación con los otros dos tipos de estancias. Este fenómeno se explica porque en muchas ocasiones el resto de estancias becadas y no becadas, sobre todo universitarias, es de un mes de duración en verano y no se produce en un contexto de inmersión en el sistema educativo del país de acogida sino que los estudiantes se inscriben en un curso específico para aprender la lengua extranjera. Esta constatación se revela de gran interés para seguir fomentando los intercambios estudiantiles en el seno de la Unión Europea, ya que hasta entonces no eran tan frecuentes las becas en otros lugares del mundo.

- Hábitos de lectura:

La influencia de los hábitos de lectura queda patente sobre todo en el léxico en español ya que en lengua extranjera los estudiantes que afirman leer algún libro al año están realizando las especialidades en Lengua Extranjera. Por tanto, la influencia sobre el léxico en lengua extranjera vendría dada por los estudios previos. Como se ha comentado anteriormente, los hábitos de lectura se relacionan de forma directa con la clase social. Se supone que si en informantes con un grado de instrucción universitario esta variable es transcendente, en hablantes con mayores diferencias sociales y culturales la incidencia de esta variable será mayor.

- Clase social:

Esta variable da cuenta de los estudios y profesiones de los padres de los informantes, así como de sus ingresos mensuales. Se confirma la hipótesis de que la influencia de esta variable está suavizada por el nivel de estudios alcanzados por los estudiantes que, en muchas ocasiones, viene marcado previamente por la clase social. La relación bivalente de la clase social con las variables lingüísticas es significativa pero no llega a marcar diferencias considerables en su comportamiento. En cambio, en el modelo causal la incidencia a través de otros factores, como los hábitos de lectura o la realización de estancias en el extranjero, cobra mayor importancia.

- Especialidad del máster:

Se ha mostrado la relación entre la especialidad del máster con los estudios previos del informante ya que los alumnos solo pueden matricularse en aquellas especialidades que guardan una relación de compatibilidad con sus estudios universitarios.³¹⁷ En caso contrario, el estudiante tiene la oportunidad de realizar un examen sobre el temario específico empleado en un concurso de oposición para esa especialidad y cuerpo.

El comportamiento de la especialidad del máster viene supeditado por el tipo de prueba lingüística. Las especialidades de lenguas extranjeras tienen una influencia positiva, sobre todo en las pruebas en inglés y francés, aunque también ayudan a fomentar el caudal léxico en español. El factor que hace de enlace entre las

³¹⁷ Orden ECI/3858/2007, de 27 de diciembre, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habilita para el ejercicio de las profesiones de profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas.

especialidades en lengua extranjera y el caudal léxico en español sería el conocimiento de otros idiomas, ya que tiene una influencia directa sobre el tipo de especialidad y sobre el léxico en lengua española.

La especialidad que muestra resultados más bajos en español y en lengua extranjera es Matemáticas. El carácter técnico de este tipo de estudios conlleva que no se fomente la competencia lingüística de la misma forma que en estudios provenientes de otras áreas. En cambio, se ha demostrado que la especialidad de Tecnología, Informática y Procesos Industriales tiene una influencia positiva sobre el léxico especializado en lengua extranjera. Entre las causas que fundamentan estos resultados podrían citarse la facilidad para identificar las palabras técnicas relacionadas con el centro de interés referido a *Ordenadores e internet* y a la exposición en sus estudios a la lengua extranjera de forma pasiva.

7.1.2. Ámbito de la psicolingüística

En este ámbito se destaca la propuesta de una taxonomía de los centros de interés. La clasificación de los centros surge a la par de los propios estudios de disponibilidad léxica. Gougenheim *et al.* (1964: 189-190) son los primeros en establecer una ordenación de los centros según el número de vocablos con objeto de discernir cuáles son los centros con mayor relevancia en los estudios de disponibilidad léxica y hacer un análisis ulterior de los mismos. El esbozo de la taxonomía expuesta en este trabajo tiene un doble objetivo. Por un lado, demostrar la inestabilidad de los resultados comparativos de estudios que cuentan con estímulos diferentes. Por otro lado, conocer la influencia de las relaciones entre las palabras según el tipo de centro de interés.

A la luz de los resultados generales en los tres idiomas, queda patente la necesidad de diferenciar los resultados según los centros de interés debido a su diferente ontogenia. Quizás si se comparara la media de palabras por informante y centro de interés de los nueve centros de este estudio con otra investigación que incluyese los dieciséis tradicionales, la media de los primeros sería superior debido a que se han eliminado centros de interés poco productivos como *Trabajos del campo y del jardín* o *Iluminación, calefacción y medios de airear un recinto*. Hasta ahora se ha

comparado el promedio de respuestas entre estudios con independencia de los centros de interés utilizados ya que según Bartol (2001: 223) es el único índice que neutraliza las diferencias en el número de informantes.

En este sentido, en los análisis comparativos del Capítulo 4 se han analizado los resultados según el promedio de respuestas provenientes de diferentes muestras con independencia de los centros de los que se han recogido los resultados. El objetivo ha sido aportar datos presentes en la mayoría de los estudios de disponibilidad léxica con los que se puedan conocer datos orientativos como la clara diferencia entre el léxico aportado por estudiantes preuniversitarios y postgraduados. Además, se muestra una tendencia común en los rangos de los centros según su índice de cohesión y densidad léxica en la propia investigación y en otros estudios. El centro más cohesionado es el primero, *Partes del cuerpo humano*, en los tres idiomas y en el estudio de Ávila Muñoz y Villena Ponsoda (2010).

No obstante, lo recomendable sería solamente comparar el promedio de respuestas de los estudios cuyo estímulo sea el mismo. Por ejemplo, Hernández Muñoz (2006: 314) contrasta la media en cada centro de interés de 17 muestras diferentes según los denominados centros tradicionales. En cambio, la autora afirma que no es igual la denominación *La escuela* que *La escuela: muebles y materiales*. En el caso de la muestra de estudiantes de Castilla la Mancha el estímulo ha sido *La escuela* dejando el campo léxico más abierto a posibles relaciones. En cambio, los estudios que han incluido *muebles y materiales* cierran más el campo de actuación de los informantes provocando que algunos escriban menos palabras. De hecho, Hernández Muñoz modifica la clasificación según la inclusividad del centro: si se trata de *La escuela* es un centro relacional y si se refiere a *La escuela: muebles y materiales* es un centro inclusivo.

En los resultados comparativos de cada idioma, se ha mostrado que las respuestas son similares según el campo en cuestión. Esto supondría que la media de palabras calculadas por centro e informante en cada lengua no fuera comparable con estudios que incluyan otros centros de interés ya que del estímulo depende en gran medida el producto y, por ende, el número de palabras y vocablos.

La clasificación propuesta en este estudio no pretende diferenciar de forma contundente la estructura de un centro con respecto a otro sino más bien elaborar una escala en la que cada centro pueda situarse según su prototipicidad. Para ello, toma como referencia los fundamentos teóricos presentes en los estudios de Hernández Muñoz (2006), Tomé Cornejo (2011), Sánchez-Saus Laserna (2011), López Fernández (2014) y como herramientas de análisis los resultados de las pruebas de disponibilidad léxica.

Concretamente, los principios en los que se basa la taxonomía realizada son: la cohesión del centro, la prototipicidad del núcleo y las relaciones entre las palabras. Cada característica da lugar a un índice que lo cuantifica: el índice de cohesión muestra la relación entre la media de palabras por informante y el número de vocablos; el índice de centralidad cuantifica la cantidad de vocablos nucleares con respecto al total de vocablos; el índice de conectividad mide la fuerza de las relaciones entre los vocablos al dividir el número de conexiones por el número total de palabras. El primer índice indica que cuanto mayor sea el índice de cohesión más homogénea será la muestra; el segundo indica que el núcleo estará más o menos definido; el tercero determina la fuerza de las conexiones entre las palabras³¹⁸. El compendio de esos tres índices da lugar al índice estructural que, como su nombre indica, sirve para definir la estructura de los centros de interés.

$$\text{Índice estructural} = IC + ICe + ICo$$

donde

IC es el índice de cohesión

ICe es el índice de centralidad

ICo es el índice de conectividad

³¹⁸ Con objeto que los valores mayores indiquen más homogeneidad y prototipicidad se ha restado uno de forma que los centros con un índice más elevado sean los que dispongan de relaciones más fuertes

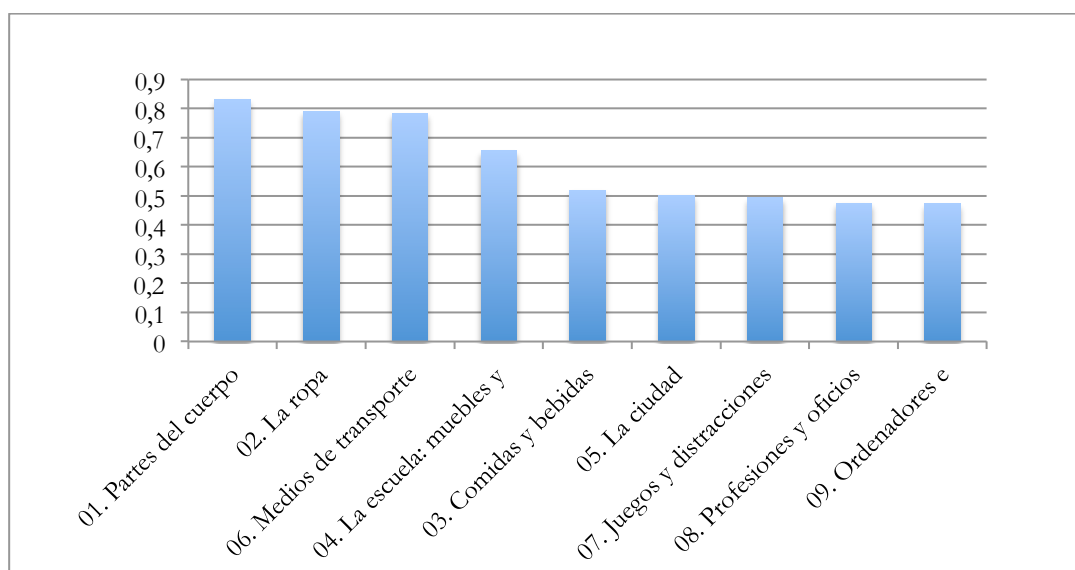


Gráfico 7.3. Clasificación de los centros de interés de este estudio según el índice estructural

Si se realizara una personificación de los centros, se podría decir que son como familias. Hay familias en las que el prototipo es más o menos numeroso y otras en las que se invita a conocidos que guardan cierta relación directa o indirecta. Cuanto menor sean los miembros de la familia, mayor comunicación y relación habrá entre ellos, como es el caso de *Partes del cuerpo humano*. En cambio, en *Medios de transporte* los vocablos o miembros de la familia son menos numerosos pero no han acudido con tanta frecuencia y cuentan con un menor número de palabras o producciones léxicas, razón por la cual se considera menos prototípico.

Al igual que ocurre en las familias, resulta complicado discernir cuál es el modelo más común. Por ello, se ha acudido a la estadística para que, igual que es posible conocer la media de hijos/as por familia, sea posible cuantificar en cada centro el léxico nuclear y diferenciarlo del resto. Para completar esta clasificación, se han analizado las principales relaciones del lexicón mental según las características de los centros de interés basados en las dos teorías que según Hernández Muñoz (2006) están presentes en los estudios de disponibilidad léxica: la teoría de los prototipos y la teoría de las redes semánticas.

Tomando como referencia la teoría de los prototipos, Ávila Muñoz y Sánchez Sáez (2010, 2011 y 2014) muestran que las palabras aportadas por los informantes giran en torno a un núcleo. En el estudio actual se observa también que las primeras palabras suelen partir del centro y se alejan de forma paulatina en forma de espiral. A

partir de un determinado momento, se vuelven a activar palabras nucleares que llevan de vuelta al núcleo, produciéndose un proceso de reentrada. A pesar de que los investigadores malagueños establecen tres modelos frecuentes con independencia del centro de interés, en la presente investigación los procesos de reentrada vienen marcados por el tipo de estímulo.

En los centros considerados más prototípicos se dan tres procesos de reentrada al núcleo: proceso de reentrada leve, moderado y pronunciado. Los informantes que aportan pocas palabras tienden a reproducir un modelo de reentrada leve en el que la mayoría de los vocablos cuentan con un elevado nivel de compatibilidad. Como el léxico nuclear está claramente definido, las palabras apenas se alejan del núcleo dibujando un proceso de reentrada leve.

A medida que los informantes cuentan con una mayor producción léxica, los procesos de reentrada se intensifican ya que no solo se aportan palabras nucleares sino también otras relacionadas de forma menos directa con el centro de interés. Por ello, los procesos de reentrada que dibujan los informantes con un número de unidades léxicas cercanas al promedio son moderados.

En el otro extremo se sitúan aquellos informantes con un caudal léxico amplio y cuyos procesos de reentrada son pronunciados. Este hecho se produce porque los hablantes no se limitan a escribir palabras nucleares sino que van más allá y mencionan palabras resultado de asociaciones individuales y otro tipo de palabras más específicas, como puede ser *hidrato* o *isquiotibial* en el centro del cuerpo humano.

En cuanto a los centros relacionales, tan solo se localizan dos modelos de reentrada. Los informantes que han aportado pocas palabras no producen modelos de reentrada leves porque el núcleo no está tan bien definido y ello provoca que no todas las palabras presenten niveles tan altos de compatibilidad. A pesar de ello, la mayoría suele pertenecer al listado de lexías nucleares.

Los informantes con un léxico en torno al promedio o superior reproducen el modelo de reentrada pronunciado con desviación lineal. Ello se debe a que los hablantes comienzan mezclando palabras muy compatibles con palabras más específicas y menos compatibles. A medida que avanzan en la secuencia léxica aportada, la reentrada al núcleo se hace con menos fuerza dibujando una línea casi

diagonal en la que tienen cabida procesos algo más moderados. En ocasiones los sujetos se alejan del núcleo pero suelen volver al activar alguna palabra concreta.

En cuanto a las conexiones entre palabras, ya en 1991 Cañizal Arévalo en su estudio realizado en Puerto Rico parte de la idea de que algunas unidades léxicas actúan como un núcleo en torno al cual se organizan las palabras en forma de constelaciones. Echeverría *et al.* (2008) secundan esta premisa crean el DispoGrafo para analizar las redes léxicas. Este programa ha permitido dibujar las redes asociativas más frecuentes entre los hablantes según el nivel de exigencia introducido (por ejemplo, se han eliminado relaciones entre palabras mencionadas por un solo informante). Asimismo, ha permitido localizar las relaciones más definidas entre los vocablos de un centro: no es igual de frecuente la relación *pizarra-tiza* que *tiza-puerta*, lo que supone un aporte más en el ámbito de la lingüística aplicada a la enseñanza de idiomas.

Al igual que ocurre con la teoría de los prototipos, en este estudio se ha comprobado que las relaciones entre palabras vienen supeditas por el tipo de centro de interés. Los centros prototípicos presentan unos subcampos léxicos mejor definidos y con mayor peso. En cambio, los centros relacionales cuentan con un mayor número de combinaciones entre palabras lo que provoca una dispersión de las relaciones.

En la misma línea que Ferreira (2006), las relaciones entre las palabras de los centros en lengua extranjera no son tan claras. No obstante, los informantes que poseen una competencia lingüística alta en lengua extranjera tienden a reproducir la secuencia léxica aportada en español. En algunos casos la poda, es decir, la selección de las relaciones según su peso, ha sido muy baja ya que en caso contrario desaparecerían todas las relaciones por no contar con una frecuencia dada.

Cabe destacar que las relaciones entre las palabras no hubiesen sido tratadas de igual forma en estudios que abogan por eliminar aquellas entradas que se desligan de los centros de interés (Samper Padilla, 1998; Carcedo González, 2001; Samper Hernández, 2003; entre otros). Por tanto, se han seguido las pautas presentes en estudios como Benítez Pérez (1992b), Mateo García (1998), Galloso Camacho (2002); Hernández Muñoz (2004), Ávila Muñoz y Villena Ponsoda (2010, 2011) que

recomiendan incluir todas las palabras con independencia de su conexión con el estímulo verbal o el número de informantes que la haya aportado ya que en caso contrario se perdería informante relevante para conocer el léxico que realmente han aportado los sujetos.

Retomando las conclusiones de los investigadores franceses, resulta curioso que en encuestas con casi medio siglo de diferencia, kilómetros de distancia y lenguas diferentes coincidan en que los centros más homogéneos en su caso o prototípicos en este estudio sean *Partes del cuerpo humano* y *La ropa*. Gracias a la elaboración de un índice estructural, es posible conocer qué tipo de configuración presenta cada centro de interés y los posibles modelos de relación léxica que se producen con el núcleo o prototipo.

7.1.3. Ámbito de la lingüística aplicada a la enseñanza de idiomas

En sintonía con el objetivo del que parten los estudios de disponibilidad, la elaboración de un francés de base, en esta investigación se ha elaborado un diccionario de léxico disponible en inglés, francés y español. Además, se ha realizado una selección léxica para crear un listado con el léxico nuclear a partir de los nueve centros de interés de la prueba y se han sentado las bases para estudios futuros.

El primer paso en la selección del léxico nuclear, se basa en la categorización conceptual de los vocablos realizada por Ávila Muñoz y Sánchez Sáez (2010, 2011) a partir de la teoría de conjuntos difusos. Esta metodología fija seis criterios de corte que establecen la pertenencia con el centro de interés: incluye los vocablos más compatibles que, a su vez, son los más disponibles. Como resultado, se obtienen siete grupos de palabras ordenados según su compatibilidad con respecto al núcleo. Los seis primeros corresponden al vocabulario más elemental y el resto al vocabulario complementario.

Para evaluar la idoneidad del vocabulario básico, se han realizado dos tipos de análisis. Por un lado, se ha contrastado el léxico nuclear en español, inglés y francés donde se ha corroborado la existencia de un prototipo en torno al cual los informantes aportan palabras muy similares. La mayoría de las palabras aportadas en lengua extranjera están presentes en los listados en español, que es la lengua con un

listado más completo por tratarse de hablantes nativos. Las palabras en lengua extranjera coinciden en más de un 80 % lo cual corrobora que independientemente del idioma, la metodología aplicada muestra el vocabulario más compatible de los informantes.

Por otro lado, el corpus se ha cotejado con las nociones específicas del Plan Curricular del Instituto Cervantes (PCIC) con las que se corresponden los centros de interés. El resultado muestra que la mayoría de las palabras del léxico básico se sitúan en los niveles A1 y A2 según el Marco Común Europeo de Referencia para las Lenguas. No obstante, no se han comparado todos los vocablos ya que el PCIC no recoge todos los sustantivos comunes ni los nombres de ciertas marcas comerciales y algunas unidades léxicas aparecen de forma parcial. Según el PCIC³¹⁹ el inventario no está cerrado sino que las palabras específicas se presentan a modo de ejemplo para ofrecer una orientación general sobre los niveles que corresponden a las unidades léxicas.

Como consecuencia, el profesor que quiera conocer si un vocablo es adecuado o no para un cierto nivel, como por ejemplo *limón*, *maíz*, *pera*, *melón*, *macarrón*, *espagueti* o *sandía* no lo encontrará en el inventario del PCIC. Además, en algunas ocasiones palabras muy compatibles son clasificadas en el PCIC en niveles de usuario independiente (B1 y B2). Es el caso de *aceite* que aparece clasificada en el nivel B1 y *refresco* que se corresponde con el nivel B2.

El léxico básico resultante de la aplicación de la teoría de los conjuntos difusos a la disponibilidad recoge todas las palabras relacionadas con una noción específica que el estudiante debe conocer en los niveles A1 y A2. La secuencia de enseñanza de los vocablos podrá tomar como criterio los seis niveles que conforman el vocabulario nuclear, de tal forma que se indica que *agua*, que pertenece al nivel 1, es más elemental que *yogur*, que se corresponde al nivel 2.

³¹⁹ «Quienes utilicen el inventario en situaciones de enseñanza y aprendizaje en las que se practique una variedad diferente habrán de hacer las adaptaciones oportunas, tanto en la selección léxica propiamente dicha como, en su caso, en la distribución por niveles. Así, por ejemplo, la serie *plátano*, *manzana*, *naranja* del subepígrafe 5.3. ‘Alimentos’ se incluye en el nivel A2 por cuanto se trata de frutas de consumo habitual en España y se considera que deben formar parte del conocimiento léxico del alumno en ese nivel» (Instituto Cervantes, 2006: 333, tomo I).

7.1.4. Propuestas y líneas de investigación futuras

Este estudio abre las puertas a investigaciones futuras de diversa índole. En primer lugar, se destaca la creación de una base de datos que recoja el léxico disponible correspondiente a las veinte nociones específicas del Plan Curricular del Instituto Cervantes. A pesar de que es un trabajo ambicioso, es un cometido viable que completa los estudios de disponibilidad con el cotejo de las palabras según el nivel del PCIC.

La disponibilidad léxica lleva asociada por antonomasia un diccionario con los listados obtenidos en las pruebas con indicación del índice de disponibilidad, la frecuencia relativa y la frecuencia acumulada. Hoy en día, las nuevas tecnologías permiten crear aplicaciones que aportan al usuario mayor información en un menor tiempo. El cometido se centra en utilizar el material del léxico disponible y crear una interfaz sencilla que permita al usuario obtener información más precisa en lapso de tiempo breve. La propuesta sería la creación de una base de datos de acceso libre que permita a los profesionales del mundo de la enseñanza obtener los siguientes datos de cada unidad léxica:

- Nivel según el MCERL. Los vocablos que no aparezcan recogidos en el DRAE se clasificarán teniendo en cuenta las relaciones con otras palabras, la clasificación de lexías similares en el PCIC y el índice de compatibilidad.
- Nombre del centro de interés donde se ha recogido la palabra y apartado concreto de la noción específica a la que hace referencia.
- Nivel de compatibilidad, índice de compatibilidad y rango en el que se sitúa con respecto al resto de vocablos de ese centro de interés.
- Índice de disponibilidad léxica y rango. El índice de disponibilidad se ve alterado por el número de informantes y por el tipo de centro de interés de forma que es aconsejable incluir el rango para poder interpretar los datos desde una perspectiva global.
- El porcentaje de aparición. Da cuenta la mayor o menor presencia de una palabra en las encuestas con independencia de su orden de aparición.

- Número total de conexiones, número de conexiones que salen y llegan a ese término en las dos direcciones. Este apartado recoge todas las relaciones diferentes que establece cada palabra en una y otra dirección. Cuanto menor sean las conexiones mayor será la fuerza de las mismas. Por ello, se indica la dirección de las mismas para introducirlas en primer lugar o a partir de otras palabras.
- Conexiones más frecuentes con indicación de la densidad. Muestra las palabras que mantienen una relación más fuerte con la lexía buscada. De esta forma, si el docente conoce que la relación entre *tiza* y *borrador* es densa, planificará en una misma secuencia didáctica la enseñanza de ambos términos.
- Enlace a la entrada del Diccionario de la Real Academia Española (DRAE). Si el usuario quiere conocer otro aspecto de un vocablo, como pueden ser sus acepciones o el género, podrá acceder al enlace que dirige directamente a esa entrada en el DRAE.

Para establecer la metodología de recogida y tratamiento de los datos de estudios futuros, se ha evaluado la adecuación de los centros de interés tradicionales a la creación de diccionarios de léxico disponible. En primer lugar, se ha analizado la literatura de referencia donde algunos mencionan que es imprescindible una revisión de los centros (Bartol Hernández, 2010; Sánchez-Saus Laserna, 2011; Paredes García, 2014; González Fernández, 2014; entre otros) con razonamientos diversos: poca productividad de los centros de interés, desfase con la vida actual, lagunas en el léxico básico, etc.

A continuación, se ha buscado una alternativa a la elección de los centros de interés que responda a las necesidades de la creación de un léxico disponible hasta llegar al Plan Curricular del Instituto Cervantes. El componente nocional del PCIC está compuesto por un inventario de *Nociones generales* y de *Nociones específicas* organizadas por temas que sirve para concretar los niveles del MCERL en español. El cuestionario de nociones específicas lleva consigo una nueva clasificación de los centros de interés creada a partir de varias obras centradas en el nivel umbral en inglés y en francés. Al igual que con los centros, supone una forma parcial de presentar la realidad dividida en temas que en ocasiones se entremezclan y en otras

pueden dejar lagunas. No obstante, el PCIC trata de aunar dentro de esas veinte nociones todos aquellos contenidos que debe adquirir un estudiante en cada uno de los niveles. Los centros de interés propuestos corresponden a las veinte nociones del PCIC (Instituto Cervantes, 2006: 331-332, tomo I):

1. Individuo: dimensión física
2. Individuo: dimensión perceptiva y anímica
3. Identidad personal
4. Relaciones personales
5. Alimentación
6. Educación
7. Trabajo
8. Ocio
9. Información y medios de comunicación
10. Vivienda
11. Servicios
12. Compras, tiendas y establecimientos
13. Salud e higiene
14. Viajes, alojamiento y transporte
15. Economía e industria
16. Ciencia y tecnología
17. Gobierno, política y sociedad
18. Actividades artísticas
19. Religión y filosofía
20. Geografía y naturaleza

Otra investigación de interés, consistiría en trasladar este mismo tipo de estudios en otras comunidades autónomas o incluso en otros países de forma que sirva para evaluar la competencia lingüística del futuro profesorado relacionada con variables propias del lugar, como el plan de estudios, la política lingüística, los requisitos de acceso a las enseñanzas, etc. Asimismo, se podría hacer énfasis en determinados aspectos cuyos resultados tengan una aplicación directa, tales como:

- El tipo de titulación del nivel B2 que acredita el futuro profesorado. Servirá para conocer en qué medida influye que la certificación esté expedida por la Escuela Oficial de Idiomas, Cambridge o Trinity en el léxico disponible.
- Un estudio pormenorizado de los principales errores en lengua extranjera.
- Vocablos ausentes en el listado del léxico nuclear para hacer más hincapié en su inclusión en los manuales.
- Evolución de otras competencias lingüísticas en comparación con la mostrada en el estudio del léxico disponible.

Se propone realizar esas comparaciones intermuestrales solo con investigaciones que tengan los mismos centros de interés. Además, se añade la propuesta de este estudio radica en comparar la media de palabras según el índice corregido de amplitud léxica (ICoAL) ya que tiene en cuenta no solo el número de palabras aportadas por un informante, sino también la relación con el centro de interés.

De esta forma, se podrá evaluar de forma más rigurosa la capacidad léxica de los informantes. No obstante, habría que establecer un consenso en las palabras consideradas nucleares (N1), asociaciones colectivas (N5), derivaciones (N25) y asociaciones individuales (N0). Cada palabra se contabiliza según la pertenencia a un grupo u otro: las asociaciones directas o palabras nucleares (N1) tienen asignados el valor de 1, las N5 un valor de 0,5, las N25 disminuye a un valor de 0,25 y por último las N0 toman un valor nulo. Esta nueva forma de contabilizar el léxico favorece la inclusión de todo tipo de palabras en los listados, sobre todo de las palabras fruto de asociaciones individuales ya que aparecerán en los últimos puestos del listado del léxico disponible, no serán contabilizadas en el ICoAL pero servirán de gran ayuda para conocer la organización de las palabras en la mente de los informantes.

En cuanto a la formación del futuro profesorado, se ha detectado que carece de formación inicial en enseñanza bilingüe. El único requisito previo consiste en la acreditación de un nivel B1 y el léxico disponible en lengua extranjera dista mucho del léxico en la lengua materna. En cambio, en Andalucía el dominio del profesorado para impartir clases bilingües tiene que ser de un mínimo de un B2 por lo que probablemente el caudal léxico sea mayor. Además, la mayoría de los informantes

desconoce conceptos básicos del ámbito de la enseñanza bilingüe, tales como, AICLE (Aprendizaje Integrado de Contenidos en Lengua Extranjera).

Se propone la creación de una materia optativa en las áreas no lingüísticas denominada enseñanza bilingüe. Según la demanda de estudiantes y la capacidad organizativa de cada universidad se pueden implantar una de estas cuatro opciones, ordenadas por orden de viabilidad. Cabe destacar que la formación general puede ser compatible con la opción de la formación según especialidad o venir incluida en forma de módulo:

1. Formación general en enseñanza bilingüe donde se haga una revisión teórica sobre los conceptos y metodologías, tales como el Aprendizaje Integrado de Contenidos en Lengua Extranjera (AICLE) y explicar en qué consiste el Plan de Fomento del Plurilingüismo. En la práctica, se propone la realización una programación que incluya unidades didácticas en lengua extranjera.
2. Formación específica según idioma. Como requisito previo para poder elegir esta asignatura los estudiantes deben acreditar un nivel B2 en lengua extranjera. Las clases se imparten en esa lengua extranjera al igual que las prácticas de los estudiantes.
3. Formación específica según especialidad. La formación inicial tiene en cuenta la materia que el estudiante pretende impartir por lo que el aprendizaje está más relacionado con la futura práctica docente. La clase se imparte en lengua materna y los informantes pueden elegir el idioma para la realización de trabajos.
- 4) Formación específica según idioma y especialidad.

Cada situación presenta sus ventajas e inconvenientes. Quizás la propuesta más sencilla consista en crear dos materias denominadas Enseñanza bilingüe (inglés) y Enseñanza bilingüe (francés) en la que se imparta un módulo con nociones generales y el futuro profesorado aprenda a programar en lengua extranjera.

En cuanto al ámbito de la psicolingüística, la disponibilidad léxica podría ser estudiada de forma transversal a través de la lingüística computacional. De esta forma, se podría crear un modelamiento computacional que permita explicar cognitivamente el orden de las palabras con numerosas aplicaciones prácticas. Es el

caso de las pruebas de fluidez verbal que son ampliamente empleadas en el ámbito de la neurolingüística y psicolingüística y, que ayuda, entre otros aspectos, a detectar patologías neurológicas.

En síntesis, las conclusiones extraídas en este estudio dan paso a líneas futuras de investigación que contribuyen a continuar el ciclo de la aplicación de un método científico. En este sentido, el químico inglés Humphry Davy (362: 1940) sugiere que:

For every discovery opens a new field for investigation of facts, shows us the imperfection of our theories. It has justly been said, that the greater the circle of light, the greater the boundary of darkness by which it is surrounded.

7.2. Conclusions

De la même façon que les objectifs et les hypothèses, les conclusions se présentent selon le cadre d'application. Parfois, il est difficile de discerner ce qui sépare chaque discipline puisqu'elles maintiennent une relation symbiotique où chacune complète l'autre. Les différentes réponses du vocabulaire en espagnol et en langue étrangère nous permettent de tirer des conclusions qui ont une application directe dans les programmes éducatifs. De plus, si l'intérêt réside dans la connaissance des différentes variables sociales, il serait dans le domaine de la sociolinguistique. Finalement, ce chapitre comprend quelques propositions et lignes de recherche futures à partir des résultats de ce travail.

7.2.1. Cadre de la sociolinguistique

Dans le domaine de la sociolinguistique, l'objectif consiste à connaître quels sont les facteurs qui ont une incidence sur le vocabulaire des informateurs et plus précisément dans le vocabulaire actif et passif. D'un côté, le vocabulaire actif se rapporte au sujet qu'il est capable d'utiliser en contexte de manière adéquate et, dans cette étude, il est obtenu par le biais de l'épreuve du lexique disponible. D'un autre côté, le vocabulaire passif peut être reconnu par le locuteur et, dans notre étude, il est associé au test d'identification de termes techniques.

D'abord, on présentera les résultats généraux d'une recherche exploratoire sur l'étude du lexique actif à partir de deux perspectives complémentaires : dans les différentes langues de l'échantillon et à partir de la comparaison avec différentes études. Ensuite, ces données se complètent avec les conclusions d'un examen statistique qui évalue l'incidence des variables sociologiques et liées à l'apprentissage d'une langue sur le lexique actif et passif de cette étude en espagnol et en langue étrangère.

7.2.1.1. Analyse générale

Les recherches du lexique disponible en espagnol, anglais et français des étudiants du Master en Professorat de l'Enseignement montrent une considérable différence entre le lexique disponible en langue maternelle et en langue étrangère. En espagnol, les informateurs ont écrit en moyenne 191,91 mots, en anglais 131,79 et en français 147,71. Les comparaisons avec les études préalables corroborent, outre que l'hypothèse du plus grand lexique disponible en langue maternelle, l'influence du niveau d'études dans le lexique disponible et dans la correcte transcription orthographique.

En espagnol comme langue étrangère, les étudiants finlandais (Carcedo González, 2000c) ont fourni une moyenne de 13,68 mots par centre d'intérêt ; les étrangers qui réalisent un séjour à Salamanque (Samper Hernández, 2002) surpassent cette moyenne avec 15,53 mots. Par contre, les espagnols de l'étude actuelle surpassent cette moyenne de 17,24 en anglais et en français. Par conséquent, le niveau atteint en langue étrangère par les étudiants espagnols est considérablement plus grand comparé aux autres cas. Parmi les raisons de cette différence entre la productivité lexicale, on peut signaler le degré de formation puisque les étudiants de cette étude ont dû certifier un minimum d'un B1 et détenir un diplôme universitaire tandis que le niveau en compétence linguistique du reste des sondés peut être inférieur.

Des résultats similaires s'obtiennent dans les comparaisons en anglais. La recherche menée par Ferreira (2006) analyse le lexique apporté par des étudiants chiliens de Pédagogie en Anglais et Traduction en Langue Étrangère et des écoliers anglais. La moyenne des étudiants natifs est de 27,28 mots, par les étudiants chiliens 18,69 et les étudiants de troisième cycle universitaire 23,19³²⁰. En outre, les écoliers écossais de la recherche de Dimitrijević (1969) ont écrit une moyenne de 20,04 mots en anglais et, par contre, les étudiants de troisième cycle universitaire dans cette langue une moyenne de 17,69 mots. Par conséquent, le facteur qui garde le rapport le plus étroit avec le lexique disponible est la compétence linguistique dans la langue de

³²⁰ La comparaison avec cette étude, de même que les précédentes, s'est réalisée à partir de la moyenne des mots dans les centres d'intérêts en commun.

l'épreuve parce que les étudiants natifs en anglais ont actualisé une majeure quantité de mots.

L'incidence directe du niveau éducatif dans la moyenne de mots s'analyse à travers la confrontation des résultats d'un échantillon d'étudiants propédeutiques (Ávila Muñoz, 2006), de population adulte stratifiée (Ávila Muñoz et Villena Ponsoda, 2010) et les étudiants de troisième cycle universitaire. La moyenne de mots par centre d'intérêt calculée sur les huit centres communs des étudiants de baccalauréat est de 19,69 mots, de 20,50 des adultes et 21,59 des étudiants de cette étude. Ces résultats vont ensemble avec les obtenus préalablement où apparaissait que les universitaires chiliens présentaient un lexique disponible en anglais plus bas que les étudiants de troisième cycle universitaire espagnols.

Par rapport à la correction orthographique, on constate que les universitaires de troisième cycle ont commis moins d'erreurs que les étudiants de baccalauréat. Les étudiants de baccalauréat madrilènes (Paredes García, 1999) commettent une moyenne de 2,26 fautes d'orthographe par sujet et centre d'intérêt et ceux aragonais 1,83 (Saura Rami, 2008) tandis que les étudiants de cette étude commentent seulement 0,28 en espagnol et 0,73 en langue étrangère. En ce qui concerne le pourcentage d'erreurs d'universitaires malaguènes (Ávila Muñoz, 2007), il est remarquablement supérieur à ceux des étudiants de troisième cycle. Par exemple, dans le premier cas il existe seulement un 6 % de sujets qui présentent aucune faute d'orthographe face au 21,05 % des étudiants de troisième cycle en espagnol. Par contre, en langue étrangère ce pourcentage est plus bas et descend jusqu'à 3,51 %.

L'analyse de la situation orthographique a permis de détecter quelles sont les erreurs orthographiques les plus usuelles afin de tirer des conclusions didactiques et critiques et de consolider les résultats des études antérieures (Paredes García, 1999; Ávila Muñoz, 2007 et Saura Rami, 2008). Les résultats révèlent que le total des fautes d'orthographe, correspondant à 66,16 %, sont dus à une inadéquade accentuation des mots. Ensuite, on trouve l'usage incorrect des graphèmes *b* et *v* puisque la poligraphie du phonème /b/ provoque des confusions, de même que l'écriture différente du phonème /x/.

Quant aux unités lexicales les plus disponibles, les vingt premiers vocables d'un même centre d'intérêt ont été analysés dans différentes études. En général, il a été constaté l'existence d'une majeure similitude entre les listages d'études qui évaluent la disponibilité d'une langue étrangère par rapport aux études en langue maternelle. De cette façon, entre les vocables les plus disponibles en langue maternelle, les sujets en ont mentionné quatre liés aux sous-vêtements tandis qu'en langue étrangère, ils en ont seulement écrit un ou deux. Ce type de vêtements est utilisé tous les jours mais ils n'apparaissent pas si fréquemment dans les manuels d'anglais, français ou espagnol comme langue étrangère.

De plus, l'espace temporaire de la pétition des données montre la diachronie de la langue : raison pour laquelle dans l'étude pionnière de Gougenheim *et al.* (1964) apparaissent des vocables, tels que *combinaison*, qui probablement de nos jours ne seraient pas si disponibles sans aucun doute parce que la façon de s'habiller a changé. C'est pour cela qu'il est possible d'observer la relation parmi l'usage et le contexte d'apprentissage dans le lexique disponible qui sera un facteur déterminant dans la sélection du lexique avec des objectifs éducatifs.

7.2.1.2. Analyse sociolinguistique

La variable utilisée pour mesurer la capacité lexicale a été l'indice corrigé de décentralisation (ICoDD). Le but de cette décentralisation lexicale est de connaître la spécificité des mots apportés par un sondé par rapport aux mots écrits par le reste de sujets. Afin d'éviter que cet indice attribue des valeurs supérieures aux locuteurs qui ont mentionné des mots fruits d'associations individuelles, les vocables se divisent entre quatre types selon la relation avec le noyau ou prototype du centre d'intérêt qui correspond au classement mené dans l'indice corrigé d'amplitude lexicale.

L'indice corrigé de décentralisation lexicale se corrèle de manière significative avec les variables qui font référence à la somme de tous les mots apportés. Dans ce cas, il est possible de distinguer l'indice d'amplitude lexicale qui prend la valeur de 1 à tous les vocables et l'indice corrigé d'amplitude lexicale qui tient en compte la relation du mot avec le centre d'intérêt et distingue quatre catégories de mots (N1, N5, N25 y N0). C'est pourquoi l'ICoDD sert à mesurer la capacité lexicale des informateurs

selon l'importance des mots au sein du centre d'intérêt et son caractère restreint. Il est probable que les locuteurs qui sont capables d'écrire les mots appartenant à un domaine spécialisé, tel que *calèche* au centre *Moyens de transport*, seront également capables d'écrire de mots partagés par la plupart de sujets, comme *voiture*.

En ce qui concerne la capacité d'identification du lexique spécialisé, celle-ci est obtenue à partir d'une épreuve qui consiste à lier des mots techniques ou spécialisés avec un des neuf centres d'intérêt qui font partie de cette étude. La variable qui en résulte tient compte de la valeur hasardeuse et intrinsèque dans toutes les épreuves de choix multiple et attribue une valeur différenciée aux réponses correctes et incorrectes. Ensuite, les réponses incorrectes pondérées se soustraient aux réponses correctes³²¹.

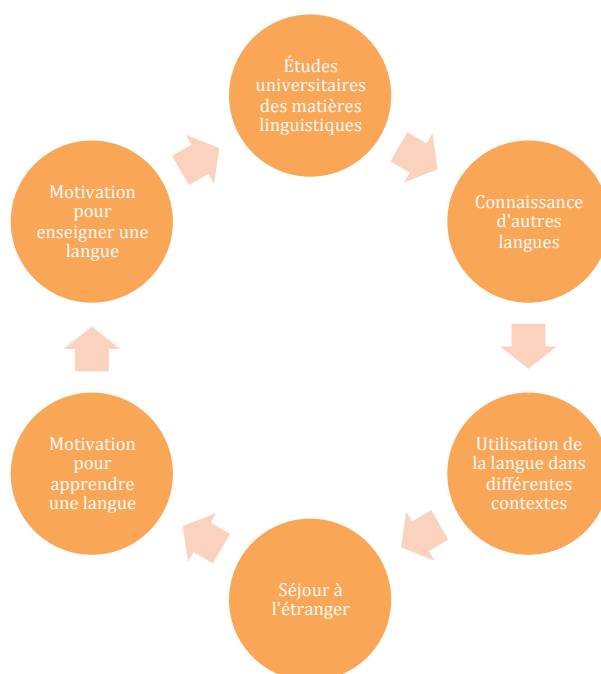
Dans l'ICoDD en langue étrangère, la variable avec la plus grande importance dans le modèle de régression linéale et causale est la spécialité du master. Les spécialités avec une influence positive dans le lexique disponible sont Langue Étrangère Anglais, Français, Langue Castellane et Littérature, Latin et Grec. Les spécialités qui maintiennent la moyenne sont Géographie et Histoire ; Dessin, Image et Arts Plastiques ; Musique ; Formation et Orientation du Travail ; Mathématiques ; Orientation Éducative ; Hôtellerie et Tourisme ; Technologie, Informatique et Procédés Industriels.

Tout d'abord, par ordre d'importance, se situe en première position le lexique disponible en espagnol qui garde une relation positive avec celui en langue étrangère. Les résultats en langue maternelle ont une incidence significative sur le lexique en langue étrangère et vice versa. Ensuite, il convient de remarquer l'incidence directe sur l'ICoDD en langue étrangère des variables qui influencent de façon directe sur l'ICoDD en langue espagnole et vice versa.

Le degré de motivation des futurs professeurs est un élément clé non seulement parce que cela influence directement sur le lexique disponible en langue étrangère, mais aussi parce qu'il est lié aux autres variables, telles que l'intérêt sur l'apprentissage d'autres langues, l'intention de se déplacer à l'étranger, une majeure utilisation de la langue, la spécialité du master, entre autres. L'usage de la langue

³²¹ Cf. section 2.5.4 pour ultérieure information sur le calcul de la variable.

étrangère dans différents contextes est le résultat d'une analyse en composantes principales (ACP) qui recueille l'information en commun des variables liées à la réalisation d'un séjour Erasmus, à l'assiduité de l'utilisation de la langue dans un domaine académique, du travail et familial.

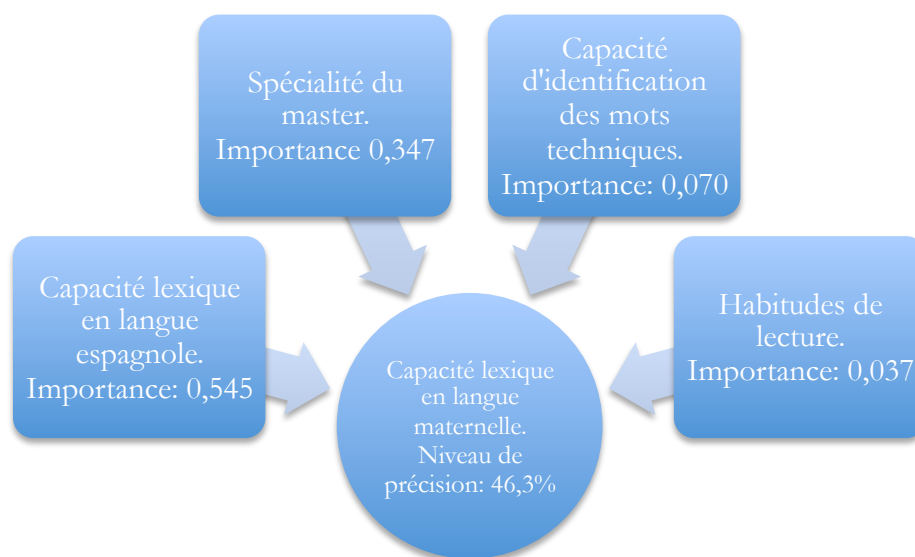


Graphique 7.1. Résumé des variables avec une importante incidence sur le lexique en langue étrangère

Les modèles multicausaux en langue espagnole expliquent avec une précision moins importante l'incidence des variables sur le lexique disponible. Il est supposé que le facteur le plus influant est le niveau de formation atteint par les sujets, qui est conditionné de façon secondaire par les autres variables. Les deux variables avec la plus grande incidence coïncident avec celles mentionnées dans le modèle en langue étrangère : le lexique en langue étrangère et la spécialité du master. De cette façon, on constate que les sujets qui présentent une meilleure capacité lexicale dans une langue la transmettent à l'autre langue grâce à la flexibilité cognitive des locuteurs bilingues.

Le comportement des spécialités du Master varie en fonction de la langue étrangère : Mathématiques est la seule spécialité qui a une influence négative ; Langue Étrangère Anglais et Français, Langue Castillane et Littérature ; Latin et Grec ; et Musique ont une incidence positive et le groupe formé par les spécialités de Géographie et Histoire ; Dessin, Image et Arts Plastiques ; Orientation Éducative ; Formation et Orientation du Travail ; Hôtellerie et Tourisme ; Technologie, Informatique et Procédés Industriels se maintiennent dans la moyenne.

La capacité d'identification de mots techniques se présente comme une variable qui favorise la capacité lexicale. Ceci explique que le bagage du vocabulaire passif influence sur le vocabulaire actif. Ce vocabulaire s'acquière grâce à la lecture qui a une répercussion mineure sur le modèle mais importante parce qu'elle peut marquer la différence entre le lexique disponible des universitaires de troisième cycle.



Graphique 7.2. Résumé du modèle de régression linéaire du lexique disponible en langue espagnole

Par rapport aux variables qui ont une incidence sur l'identification du lexique spécialisé, il y a deux variables qui coïncident en langue étrangère et en langue espagnole : l'usage de la langue dans différents contextes et le type d'épreuve. Ceci entraîne que les locuteurs avec une meilleure maîtrise des mots techniques en langue espagnole, seront aussi capables d'identifier le lexique spécialisé en langue étrangère plus facilement. Et en plus, l'indice corrigé d'identification de mots techniques (ICIMT), garde une relation significative avec la capacité lexicale et avec la langue de l'épreuve. En anglais, les sujets ont fait descendre la moyenne des résultats tandis qu'en français les sondés l'ont fait augmenter légèrement. Ainsi, la langue choisie a une certaine répercussion sur les résultats finaux de l'épreuve.

Finalement, la classe sociale influence de manière directe et indirecte sur la capacité lexicale. Dans le modèle causal, on a recueilli l'influence à travers de la variable de l'usage de la langue et de la capacité lexicale en espagnol. En outre, les spécialités qui présentent les plus grandes influences sont Langue Étrangère Anglais et Français. Parallèlement, la spécialité en Langue Étrangère est liée au nombre de

langues que parle un sujet et au niveau de compétence linguistique ainsi qu'à l'usage de la langue étrangère dans différents contextes, entre autres.

Les résultats en espagnol montrent une majeure connaissance du lexique spécialisé qu'en langue étrangère. La spécialité présente une incidence différente. Dans ce cas, les variables avec une influence positive sont Langue Étrangère Français et Technologie, Informatique et Procédés Industriels. La seule spécialité qui fait descendre l'indice corrigé d'identification des mots techniques (ICMT) est Mathématiques. Le reste des spécialités (Langue Étrangère Anglais ; Langue Castellane et Littérature ; Latin et Grec ; Dessin, Image et Arts Plastiques ; Orientation Éducative, Formation et Orientation du Travail ; Hôtellerie et Tourisme ; Géographie et Histoire) se maintient autour de la moyenne.

En langue étrangère, la capacité lexicale a aussi une influence directe. Le lexique spécialisé en langue étrangère se montre remarquable puisque la connaissance d'autres langues a une incidence non seulement sur les épreuves en langue mais elle aide aussi à développer la maîtrise linguistique dans la langue maternelle. Les variables qui montrent les habitudes de lecture et l'exposition aux programmes culturels présentent une influence indirecte à travers de la classe sociale et directe sur le lexique. Cet aspect a une grande répercussion didactique puisqu'il suggère, d'une part, que les habitudes de lecture ne sont pas les mêmes dans toutes les classes sociales et d'autre part, il aide à améliorer la compétence lexicale des étudiants. L'influence de la classe sociale dans le modèle causal dans l'ICILS en espagnol est plus important que dans le reste de variables dépendantes (ICILS en langue étrangères et ICoDD en espagnol et en langue étrangère).

En résumé, on présentera les variables les plus importantes dans le lexique actif et passif qui ont répondu aux hypothèses de cette étude. Il faut souligner que celles qui n'aident pas à expliquer le comportement du lexique, telles que le sexe, ne sont pas incluses. De même que l'étude d'Ávila Muñoz et Villena Ponsoda (2010 : 285), le sexe n'est pas une variable importante étant donné que les inégalités dans les rôles masculins et féminins disparaissent au fur et à mesure qu'augmente la situation socioéconomique des personnes.

- Degré de formation :

Par rapport aux enquêtes de disponibilité lexicale d'étudiants de baccalauréat, une augmentation considérable dans le nombre de mots apportés par les étudiants du troisième cycle universitaire est observée. Ce décalage d'instruction diminue l'importance de variables présentes dans d'autres recherches, telles que le type de centre éducatif, puisque les études spécialisées font disparaître ces différences.

- Connaissance d'autres langues :

Les données apportées par les sujets qui parlent plus d'une langue étrangère montrent une meilleure capacité lexicale en espagnol et, notamment, en langue étrangère. Il paraît indispensable de réaliser une analyse plus détaillée sur la compétence linguistique qu'ont atteint les locuteurs dans chaque langue étrangère puisque celles-ci ont une incidence sur les résultats. De cette façon, les sondés qui connaissent une langue avec un niveau B2 disposent d'une plus grande capacité lexicale que ceux qui connaissent deux langues avec un niveau débutant et moyen car généralement ce niveau B2 correspond à la langue de l'épreuve (soit anglais ou français).

- Séjours à l'étranger :

Les séjours se divisent selon leurs durées et leurs financements en se distinguant entre le séjour Erasmus, d'autres séjours avec une aide boursière et d'autres séjours sans bourse. Sans aucun doute, le programme Erasmus montre une plus grande relation avec la compétence lexicale et le lexique spécialisé en langue étrangère en comparaison avec les autres deux types de séjour. Ce phénomène s'explique parce que dans beaucoup d'occasions le reste des séjours boursiers, surtout les universitaires, sont d'une durée d'un mois en été et correspondent à un cours spécifique pour apprendre la langue étrangère. Cette constatation se révèle de grand intérêt pour favoriser l'encouragement des échanges universitaires au sein de l'Union Européenne et dans d'autres pays où les étudiants apprennent une langue dans un contexte d'immersion au système éducatif du pays d'accueil.

- Habitudes de lecture :

L'influence des habitudes de lecture se manifeste surtout sur le lexique en espagnol puisqu'en langue étrangère les étudiants, qui assurent lire quelques livres par an, font des études de spécialités en Langue Étrangère. De cette manière, l'incidence

sur le lexique en langue étrangère est liée aux études préalables. Tel qu'il est indiqué auparavant, les habitudes de lectures gardent un rapport direct avec la classe sociale. On suppose que si l'influence de cette variable sur des sujets avec un niveau de formation universitaire est considérable, sur des locuteurs qui présentent des différences sociales et culturelles plus importantes, alors l'incidence sera plus grande.

- Classe sociale :

Cette variable tient compte des études, professions des parents des informateurs et ses revenus mensuels. On ratifie l'hypothèse que l'influence de cette variable est adoucie par le niveau d'études atteintes par les étudiants qui, dans beaucoup d'occasions est marqué préalablement par la classe sociale. La relation bivariée de la classe social avec les variables linguistiques est significative mais cela n'arrive pas à marquer des différences considérables dans leurs comportements. Néanmoins, dans le modèle causal l'incidence à travers d'autres facteurs, tels que les habitudes de lecture ou la réalisation de séjours à l'étranger donne à la classe sociale une plus grande importance.

- Spécialité du master :

Il a été démontré auparavant le lien entre la spécialité du master avec les études précédentes du sujet : les universitaires peuvent s'inscrire seulement dans les spécialités ayant une relation directe avec leurs études antérieures. Dans le cas contraire, l'étudiant a l'opportunité de passer un examen sur le programme spécifique de la matière et corps du concours.

Le comportement de la spécialité du master dépend du type d'épreuve linguistique. Les spécialités de langues étrangères ont une influence positive, surtout dans les épreuves en anglais et en français, même si celles-ci favorisent aussi la capacité lexicale en espagnol. Le facteur qui lie les spécialités en langue étrangère et le lexique disponible en espagnol serait la connaissance d'autres langues puisqu'elle a une influence directe sur le type de spécialité et sur le lexique en langue espagnole.

La spécialité qui montre les résultats les plus bas en espagnol et en langue étrangère est Mathématiques. Le caractère technique de ce type d'études conduit à un encouragement différent de la compétence linguistique en comparaison avec les études dans d'autres domaines. Cependant, la spécialité Technologie, Informatique et

Procédés Industriels a une influence positive sur le lexique spécialisé en langue étrangère. Cela peut s'expliquer grâce à la majeure identification des termes techniques liée au centre d'intérêt *Ordinateurs et Internet* et à l'exposition passive des études sur la langue étrangère.

7.2.2. Cadre de la psycholinguistique

Dans ce cadre il est remarquable l'ébauche d'une taxonomie des centres d'intérêt. Le classement des centres apparaît au même temps que les propres études de disponibilité lexicale. Gougenheim *et al.* (1964 : 189-190) sont les premiers qui établissent un ordre selon le nombre de vocables afin de discerner quels sont les centres les plus importants pour les études de disponibilité lexicale et faire une analyse ultérieure. La proposition de l'organisation de cette étude a un double objectif. D'un côté, elle prétend démontrer l'instabilité des résultats comparatifs des études avec différents centres. D'un autre côté, elle vise à connaître l'influence des relations entre les mots selon le type de centre d'intérêt.

D'après les résultats généraux dans les trois langues, il est nécessaire de différencier les données selon les centres d'intérêt dû à leurs différentes ontogénies. Peut-être que si l'on fait la comparaison de la moyenne des mots par sujet et centre d'intérêt des neuf centres de cette étude avec une autre recherche qui comprend les seize centres traditionnels, la première serait supérieure parce que celle-ci n'inclut pas les centre d'intérêt peu productifs, tels que *Les travaux des champs et du jardinage* ou *Le chauffage et l'éclairage*. Pour l'instant, nombreuses sont les études qui ont comparé la moyenne de réponses entre des échantillons indépendamment des centres d'intérêt utilisés puisque selon Bartol Hernández (2001 : 223) est le seul indice qui neutralise les différences entre le nombre des sondés.

En ce sens, les analyses comparatives du Chapitre 4 ont analysé les résultats selon la moyenne des réponses de différents échantillons sans tenir compte des centres d'intérêt de chaque étude. L'objectif a consisté à apporter des données présentes dans la plupart des études de disponibilité lexicale avec lesquels on peut faire apparaître des résultats, tels que la différence entre le lexique des étudiants de baccalauréat et les universitaires de troisième cycle. En outre, ce chapitre-là démontre

une tendance commune de tous les rangs des centres selon leur indice de cohésion et densité lexicale dans la propre recherche et dans d'autres études. Le centre qui montre le plus de cohésion est le premier, *Les parties du corps*, dans les trois langues et dans l'étude d'Ávila Muñoz et Villena Ponsoda (2010), entre autres.

Il conviendrait de comparer seulement la moyenne des réponses des études qui incluent les mêmes centres d'intérêt. Par exemple, Hernández Muñoz (2006: 314) compare la moyenne de chaque centre d'intérêt de dix-sept échantillons différents selon les dénommés centres traditionnels. Néanmoins, la chercheuse affirme que la dénomination de *L'école* n'est pas la même que *L'école, ses meubles et son matériel scolaire*. Dans l'étude de Castille le centre utilisé est *L'école* qui laisse le champ lexicale plus ouvert à de possibles relations. Par contre, les études qui ont inclus *L'école : ses meubles et son matériel scolaire* ferment plus le champ d'actuation des sondés et par conséquent, ils ont écrit un nombre inférieur de mots. De fait, Hernández Muñoz change le classement selon l'inclusivité du centre : s'il s'agit que *L'école* est un centre relationnel mais dans le cas de *L'école, ses meubles et son matériel scolaire*, il s'agit d'un centre inclusif.

Les résultats comparatifs de chaque langue ont montré que les réponses sont similaires selon le centre d'intérêt concret. Cela suppose que la moyenne de mots calculée par centre et par sujet dans chaque langue ne soit pas comparable avec les études qui incluent d'autres centres d'intérêt puisque ceci modifie, en grande partie, les réponses et, par conséquent, le nombre de mots et de vocables.

Le classement proposé dans cette étude ne prétend pas différencier de façon catégorique la structure d'un centre par rapport à un autre, il vise plutôt à élaborer une échelle où chaque centre puisse se situer selon sa prototypicité. Pour cela, il fait allusion aux fondements théoriques présents dans les études de Hernández Muñoz (2006), Tomé Cornejo (2011), Sánchez-Saus Laserna (2011), López Fernández (2014) et comme outil d'analyse des résultats des épreuves de disponibilité lexicale.

Les principes de base sont les suivants : cohésion du centre, la prototypicité du noyau et les relations entre les mots. Chaque caractéristique donne lieu à l'indice qui le quantifie : l'indice de cohésion montre la relation entre la moyenne de mots par sujet et le nombre de vocables ; l'indice de centralité mesure la quantité de vocables

nucléaires par rapport au total de vocables ; l'indice de connectivité mesure la force des relations entre les vocables en divisant le nombre de connexions par le nombre total de mots. Dans le premier cas, plus l'indice de cohésion est plus haut, plus homogène est l'échantillon. Le second indique que le noyau sera plus ou moins défini et le troisième détermine la force des connexions entre les mots. L'addition de ces trois indices donne lieu à l'indice structural qui, comme sa propre dénomination l'indique, sert à définir la structure des centres d'intérêt.

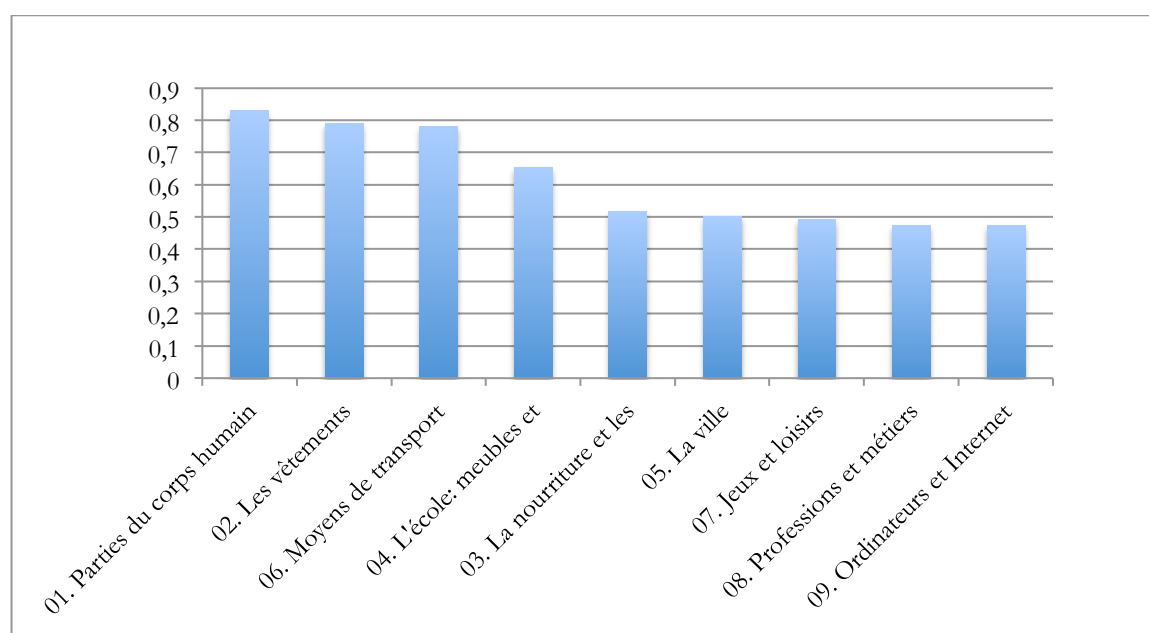
$$\text{Indice structural} = IC + ICe + ICo$$

où

IC est l'indice de cohésion

ICe est l'indice de centralité

ICo est l'indice de connectivité



Graphique 7.3. Classement des centres d'intérêt de cette étude selon l'indice structural

Si l'on réalise une personnification des centres, on pourrait dire qu'ils sont comme des familles. Il y a des familles où le prototype est plus ou moins nombreux et d'autres où l'on invite des amis qui gardent une certaine relation directe ou indirecte. Moins il y a des membres de la famille, plus de communication et de relation il y aura entre eux, tels que le centre *Les parties du corps*. Par contre, dans le centre *Moyens de transport* les vocables (ou les membres de la famille) sont moins

nombreux mais ils ne sont pas apparus si fréquemment et ce centre compte un nombre de mots plus petits, c'est pourquoi il est considéré moins prototypique.

De la même façon que dans les familles, il est complexe de discerner le modèle le plus commun mais l'on peut connaître quelle est la moyenne d'enfants par famille. Pour cette raison, on a fait appel à la statistique pour quantifier le lexique nucléaire de chaque centre et le différencier du reste. Pour compléter ce classement, on a analysé les relations principales du lexique dans l'esprit selon les caractéristiques des centres d'intérêt basés sur deux théories, qui selon Hernández Muñoz (2006), sont présentes dans les études de disponibilité lexicale : la théorie des prototypes et la théorie des réseaux sémantiques.

En ce qui concerne la théorie des prototypes, Ávila Muñoz y Sánchez Sáez (2010, 2011 y 2014) montrent que les mots apportés par les sondés tournent autour d'un noyau. Les premiers mots partent en général du noyau et s'éloignent de manière lente en spirale. À partir d'un moment donné, les mots nucléaires retournent au centre en produisant un procédé de re-entrée. Ces chercheurs ont établi trois modèles fréquents dans tous les centres d'intérêt. Cependant, dans cette étude les procédés de ré-entrée sont marqués par le type de centre.

Dans tous les centres considérés prototypiques, il se produit trois procédés de re-entrée au noyau : procédé de re-entrée léger, modéré et prononcé. Les sujets qui apportent moins de mots ont l'habitude de reproduire un modèle de re-entrée léger dans lequel la plupart des vocables ont un niveau élevé de compatibilité. Étant donnée que le lexique nucléaire est clairement défini, les mots s'éloignent à peine du noyau et dessinent un procédé de ré-entrée léger.

Au fur et à mesure que les locuteurs écrivent une plus grande quantité des mots, les procédés de re-entrée s'intensifient puisqu'ils n'apportent pas seulement de mots nucléaires sinon qu'ils incluent aussi de mots liés de manière moins directe au centre d'intérêt. Pour cette raison, les procédés de re-entrée que dessinent les sujets avec un nombre d'unités lexicales proches à la moyenne sont modérés.

À l'autre extrémité se situent les sondés avec une capacité lexicale approfondie dont les procédés de re-entrée sont prononcés. Cela se produit parce que les locuteurs ne se limitent pas qu'à écrire des mots nucléaires sinon qu'ils vont plus loin

et mentionnent des mots fruit d'associations individuelles et d'autres types de mots plus spécifiques, tels que *hydrate* ou *ischion* dans le centre d'intérêts, *Les parties du corps*.

Par rapport aux centres relationnels, ils présentent seulement deux modèles de re-entrée. Les sujets qui ont apporté peu de mots ne produisent pas de modèles de re-entrée légers parce que le noyau n'est pas bien défini et provoque ainsi un décalage du niveau de compatibilité des mots. Malgré cela, la plupart d'entre eux appartiennent au lexique nucléaire.

Les locuteurs ayant un lexique autour de la moyenne ou supérieur reproduisent le modèle de re-entrée prononcé avec une déviation linéale. Cela s'explique parce que les sujets commencent à entremêler des mots très compatibles avec des mots plus techniques et moins compatibles. À mesure que la séquence lexicale avance, la re-entrée au noyau se fait avec moins de force en dessinant une ligne presque diagonale où se produisent des procédés plus modérés. Dans certaines occasions, les sondés s'éloignent du noyau mais ils y retournent souvent à travers de l'activation d'un mot plus concret.

Quant aux connexions des mots, l'étude menée par Cañizal Arévalo (1991) à Porto Rico évoquait déjà l'idée que certaines unités lexicales remplissent la fonction de noyau autour duquel s'organisent les mots sous la forme de constellations. Echeverría *et al.* (2008) reprennent cette prémisse et créent le DispoGrafo pour analyser les réseaux lexicaux. Ce programme a permis de dessiner les réseaux associatifs plus fréquents entre les locuteurs selon le niveau d'exigence introduit (par exemple, les relations entre mots mentionnés par un seul sujet ont été éliminés). En outre, il a permis de localiser les relations plus définies entre les vocables d'un centre : il n'est pas si fréquent de trouver la relation *tableau-craie* et *craie-porte*, ce qui suppose un apport en plus dans le cadre de la linguistique appliquée à l'enseignement des langues.

Sur cette lancée de la théorie des prototypes, cette étude montre que les relations entre des mots sont indépendants du type de centre d'intérêt. Les centres prototypiques présentent des sous-champs lexicaux mieux définis et avec un poids plus lourd. Cependant, les centres relationnels possèdent un plus grand nombre de combinaisons entre les mots, ce qui provoque une dispersion des relations.

Dans ce sens, Ferreira (2006) expose que les relations entre les mots des centres d'intérêt ne sont pas si claires. Néanmoins, les sujets qui possèdent une compétence linguistique élevée en langue étrangère tendent à reproduire la séquence lexicale apportée en espagnol. Dans certains cas, la taille, c'est à dire, la sélection des relations selon leur poids, a été très faible car, dans le cas contraire, toutes les relations pourraient disparaître par manque d'une fréquence donnée.

Il faut souligner que les analyses de relations entre des mots n'auraient pas pu être de la même façon au sein des études qui plaident pour éliminer les mots qui se dissocient des centres d'intérêt (Samper Padilla, 1998; Carcedo González, 2001; Samper Hernández, 2003; entre autres). Par conséquent, on a suivi les règles présentes dans des études telles que Benítez Pérez (1992b), Mateo García (1998), Galloso Camacho (2002); Hernández Muñoz (2004) qui recommandent d'inclure tous les mots indépendamment de leur connexion avec le centre d'intérêt ou le nombre de sondés qui les ont apportés puisque l'on risque de perdre des données de grande valeur pour connaître le lexique que les sujets ont vraiment écrit.

Si l'on reprend les conclusions des chercheurs français, il est curieux de voir que dans des enquêtes avec quasi un demi siècle de différence, des kilomètres de distance et de nombreuses langues différentes, on trouve à nouveau que les centres les plus homogènes ou prototypiques de cette étude soient *Les parties du corps* et *Les vêtements*. C'est grâce à l'élaboration d'un indice structurel qu'il est possible de connaître quel type de configuration présente chaque centre d'intérêt et les possible modèles de relation lexicale avec le noyau ou le prototype.

7.2.3. Domaine de la linguistique appliquée à l'enseignement des langues

En harmonie avec le but principal des études de disponibilité lexicale, l'élaboration d'un français de base, on a élaboré dans cette étude un dictionnaire de lexique disponible en espagnol, anglais et français. Et en plus, on a sélectionné le vocabulaire pour créer un lexique nucléaire à partir des neuf centres d'intérêt puis établir les bases pour les études à venir.

Le premier pas dans la sélection du lexique nucléaire, se base sur la catégorisation conceptuelle des vocables menée par Ávila Muñoz y Sánchez Sáez (2010, 2011) à partir de la théorie des ensembles diffus. Cette méthodologie fixe six critères pour tailler les données qui établissent l'appartenance avec le centre d'intérêt : ceux-ci incluent les vocables les plus compatibles qui, au même temps, sont les plus disponibles. Comme résultat, il est possible d'obtenir sept groupes de mots ordonnés selon leur compatibilité avec le noyau. Les six premiers correspondent au vocabulaire de base et le reste au vocabulaire complémentaire.

Pour évaluer l'adéquation du vocabulaire de base, on a réalisé deux types d'analyse. D'une part, on a contrasté le lexique nucléaire en espagnol, anglais et français où on a prouvé l'existence d'un prototype autour duquel les informateurs apportent des mots très similaires. La plupart des mots apportés en langue étrangère se trouve dans les listes en espagnol, qui est la langue avec la liste la plus complète parce que les locuteurs sont natifs. Les mots en langue étrangère coïncident à hauteur de 80 %, ce qui corrobore qu'indépendamment de la langue, la méthodologie appliquée montre le vocabulaire plus compatible des sujets.

D'un autre côté, le corpus s'est confronté aux notions spécifiques du Plan Curriculaire de l'Instituto Cervantes (PCIC) (Instituto Cervantes, 2006) qui correspondent aux centres d'intérêt. Le résultat montre que la plupart des mots du lexique basique se situent sur les niveaux A1 et A2 selon le Cadre européen commun de référence pour les langues (Conseil de l'Europe, 2001). Cependant, tous les vocables n'ont pas pu être comparés parce que le PCIC n'inclut pas tous les substantifs communs ni les noms de certaines marques commerciales et quelques unités lexicales apparaissent partialement. Selon le PCIC, l'inventaire n'est pas fermé car les termes spécifient les spécifiques se présentent en guise d'exemple pour offrir une orientation générale sur les niveaux qui correspondent aux unités lexicales.

Par conséquent, le professeur qui veut savoir si un vocable est adéquat ou non pour un certain niveau, comme par exemple *citron*, *maïs*, *poire*, *melon*, *macaroni*, *spaghetti* ou *pastèque*, ne les trouvera pas dans l'inventaire du PCIC. En outre, dans quelques cas des mots très compatibles sont classifiés dans le PCIC dans de niveaux

d'utilisateur indépendant (B1 et B2). C'est le cas de *huile* qui apparaît classifiée dans le niveau B1 et *rafraîchissement* qui correspond au niveau B2.

Ceux qui utilisent l'inventaire dans des situations d'enseignement et d'apprentissage où se pratique une variété différente devront réaliser des adaptations opportunes, tantôt dans la sélection lexicale proprement dite, tantôt dans son cas, dans la distribution par niveaux. Donc, par exemple, la série *banane, pomme, orange* de la sous-épigraphie 5.3. « Aliments » s'insère dans le niveau A2 puisqu'il s'agit des fruits de consommation naturelle en Espagne est considéré comme faisant partie de la connaissance lexicale de l'élève dans ce niveau (Instituto Cervantes, 2006: 333, tome I).

Le lexique basique qui résulte de l'application de la théorie des ensembles diffuses à la disponibilité lexicale recueille tous les mots liés à la notion spécifique que l'étudiant doit connaître dans les niveaux A1 et A2. La séquence d'enseignement des vocables pourra alors avoir comme critère les six niveaux qui composent le vocabulaire nucléaire, de telle façon qu'il est possible d'indiquer que le mot *eau*, qui appartient au niveau 1, est plus élémentaire que *yaourt*, qui fait partie du niveau 2.

7.2.4. Propositions et lignes de recherche futures

Cette étude ouvre des portes aux recherches futures de différente nature. Tout d'abord, il est important de souligner la création d'une base de données ou corpus qui recueille le lexique basique correspondant aux vingt notions spécifiques du Plan Curriculaire de l'Institut Cervantès. Même si ce travail est ambitieux, c'est un objectif viable qui complète les études de disponibilité avec la confrontation des mots selon le niveau du PCIC.

La disponibilité lexicale est associée par antonomase à un dictionnaire avec les listes de vocables obtenus dans les épreuves avec indication de l'indice de disponibilité, la fréquence relative et la fréquence accumulée. De nos jours, les nouvelles technologies permettent de créer des applications qui apportent à l'utilisateur la plus grande information dans le temps le plus moindre. Le but est centré sur l'utilisation du matériel du lexique disponible et la création d'une interface simple qui permet à l'utilisateur d'obtenir une information plus précise dans un laps

de temps. La proposition serait l'élaboration d'une base de données ou d'un corpus d'accès libre qui permet aux professionnels du monde de l'enseignement d'obtenir les renseignements suivants de chaque unité lexicale :

- Niveau selon le Cadre européen commun de référence pour les langues (CECRL). Les vocables qui n'apparaissent pas dans le PCIC seront classifiés par rapport aux relations avec les autres mots, le classement de mots similaires dans le PCIC et l'indice de compatibilité.
- Nom du centre d'intérêt où apparaît le mot et section concrète de la notion spécifique à laquelle il fait référence.
- Niveau de compatibilité, indice de compatibilité et rang où se situe le mot par rapport au reste de vocables de ce centre d'intérêt.
- Indice de disponibilité lexicale et rang. L'indice de disponibilité est modifié par le nombre de sondés et par le type de centre d'intérêt de façon qu'il est convenable d'inclure le rang pour pouvoir interpréter les données à partir d'une perspective globale.
- Le pourcentage d'apparition. Ceci enregistre la présence d'un mot dans les épreuves indépendamment de son ordre d'apparition.
- Nombre total de connections, nombre de connections qui partent et qui arrivent jusqu'à ce terme. Cette section recueille toutes les relations différentes qu'établit chaque mot dans les deux directions. Moins de connections il y aura, plus grande leur force sera.
- Connections fréquentes avec indication de la densité. Ceci montre les mots qui gardent une relation plus forte avec le vocable cherché. De cette façon, le professeur sait que la relation entre *craie* et *effaceur* est dense, planifiera dans une même séquence didactique l'enseignement des deux termes.
- Lien à l'entrée du dictionnaire de la Real Academia Española. Si l'utilisateur veut connaître un autre aspect d'un vocable, tel que ses acceptions ou le genre, il pourra accéder au lien qui dirige directement à cette entrée dans le dictionnaire.

Pour établir la méthodologie de récolte et de traitement des données dans des recherches futures, on a évalué l'adéquation des centres d'intérêt traditionnels à la

création de dictionnaires de lexique disponible. En premier lieu, on a analysé la littérature de référence dans laquelle certains chercheurs proposent de faire une révision des centres (Bartol Hernández, 2010; Sánchez-Saus Laserna, 2011; Paredes García, 2014; González Fernández, 2014; entre autres) à partir de raisonnements différents : peu de productivité des centres, déphasage avec la vie actuelle, lacunes dans le lexique basique, etc.

Ensuite, on a cherché une alternative à l'élection des centres d'intérêt qui répond aux nécessités de la création d'un lexique disponible pour arriver au Plan Curriculaire de l'Institut Cervantès. Le composant notionnel du PCIC est composé par un inventaire de Notions générales et de Notions spécifiques organisées par thèmes qui sert à concrétiser les niveaux du CECRL en espagnol. L'épreuve de notions spécifiques apporte un nouveau classement des centres d'intérêt créé à partir de plusieurs ouvrages sur le niveau seuil en anglais et en français. Ceci implique, de même qu'avec les centres d'intérêt, une manière partielle de présenter la réalité fragmentée en thèmes qui dans certains cas s'entremêlent et dans d'autres cas laisse des lacunes lexicales. Néanmoins, le PCIC essaie de rassembler dans ces vingt notions tous les contenus que doit acquérir un étudiant dans chaque niveau. Les centres d'intérêt proposés correspondent aux vingt notions du PCIC (Instituto Cervantes, 2006: 331-332, tome I):

1. Individu : dimension physique
2. Individu : dimension perceptive et animique
3. Identité personnelle
4. Relations personnelles
5. Alimentation
6. Éducation
7. Travail
8. Loisir
9. Information et moyens de communication
10. Logement
11. Services
12. Achats, magasins et établissements

13. Santé et hygiène
14. Voyages, logement et transport
15. Économie et industrie
16. Science et technologie
17. Gouvernement, politique et société
18. Activités artistiques
19. Religion et philosophie
20. Géographie et nature

Une autre recherche d'intérêt, consiste à déplacer ce type d'études dans d'autres communauté autonome ou même dans d'autres pays de manière à évaluer la compétence linguistique des futurs professeurs liée aux variables propres du lieu, telle que le plan d'études, la politique linguistique, les requises d'accès à l'enseignement, etc. Sur cette lancée, on propose l'évaluation du lexique virtuel et spécialisé des professeurs qui donnent des cours de matières non linguistiques . En outre, on pourrait mettre l'accent sur des aspects divers dont les résultats auront une application directe, tels que :

- Le type de diplôme du niveau B2 qui certifie le futur professorat. Cela aidera à connaître dans quelle mesure influence l'organisme de délivrance de la certification (École Officielle des Langues, Cambridge English ou Trinity College) sur le lexique disponible.
- Une étude détaillée sur les erreurs principales en langue étrangère et en langue espagnole et les possibles circonstances.
- Mots absents dans la liste de lexique nucléaire pour mettre l'accent sur les manuels.
- Évolution d'autres compétences linguistiques en comparaison avec celle du lexique disponible.

On propose de réaliser ces comparaisons entre différentes recherches à condition qu'elles examinent les mêmes centres d'intérêt. En plus, on ajoute la proposition de cette étude qui consiste à comparer la moyenne de mots selon l'indice corrigé d'amplitude lexicale (ICoAL) puisqu'il tient compte non seulement des mots écrits par un sujet, mais aussi sa relation avec le centre d'intérêt. Ainsi, il sera possible

d'évaluer d'une façon plus rigoureuse la capacité lexicale des sondés. Cependant, il serait nécessaire d'établir un consensus entre les mots considérés nucléaires (N1), associations collectives (N5), dérivations (N25) et associations individuelles (N0). Chaque mot se comptabilise selon l'appartenance à un groupe : les associations directes ou les mots nucléaires (N1) se sont vus assigner la valeur 1, les N5 la valeur de 0,5, les N25 diminuent à 0,25 et finalement, les N0 acquièrent une valeur nulle. Cette nouvelle manière de comptabiliser le lexique favorise l'inclusion de tout type de mots dans les listes, surtout des mots fruit d'associations individuelles mais ils apparaissent dans les dernières positions des listes du lexique disponible et ne seront pas comptabilisés dans le ICoEL. Le plus important est que ces mots seront d'une grande aide pour connaître l'organisation des mots dans l'esprit des sondés.

Ces résultats pourraient se compléter avec des analyses plus concrètes et générales. D'un côté, on pourrait réaliser une étude détaillée des principales erreurs en langue étrangère et des mots manquant en comparaison avec le lexique basique pour mettre l'accent sur les manuels. D'un autre côté, on pourrait faire des recherches sur l'évolution des compétences linguistiques, telle que l'expression orale, la compréhension écrite, etc.

Quant à la formation des futures professeurs, on a détecté qu'il manquait de formation initiale dans l'enseignement bilingue. La seule requise préalable consiste à fournir une accréditation d'un niveau B1, au minimum, dans une langue étrangère parlée dans l'Union Européenne. Par contre, le lexique disponible en anglais et en français s'éloigne beaucoup du lexique en langue maternelle. En Andalousie la compétence linguistique pour donner des cours bilingues établi un minimum d'un niveau B2, raison pour laquelle la capacité lexicale des professeurs qui actuellement travaillent est probablement meilleure. En outre, la plupart des sujets ne connaissent pas de concepts basiques du domaine de l'enseignement bilingue, tels que l'apprentissage intégré du contenu et de la langue.

On propose la création d'une matière optionnelle dans les spécialités non linguistiques dénommée *enseignement bilingue*. Selon la demande d'étudiants et la capacité d'organisation de chaque université, il est possible d'implanter l'une de ces quatre options, ordonnées selon leur viabilité. Il faut souligner que la formation

générale peut être compatible avec l'option de la formation selon spécialité ou être incluse comme un module :

1. Formation générale en enseignement bilingue où on réalise un parcours théorique sur les concepts et méthodologies, tels que l'apprentissage intégré du contenu et de la langue, et expliquer en quoi consiste le *Plan de Fomento del Plurilingüismo* (Junta de Andalucía, 2005). Pour la pratique, on propose la réalisation d'un programme qui inclut des unités didactiques en langue étrangère.
2. Formation spécifique selon la langue. Comme préalablement requise pour pouvoir choisir cette matière, les étudiants devront accréditer un niveau B2 en langue étrangère. Les cours seront donnés dans la langue étrangère de même que les stages des étudiants.
3. Formation spécifique selon la spécialité. La formation initiale tient compte de la matière que l'étudiant prétend donner puisque l'apprentissage doit être plus lié à la future pratique d'enseignement. Les cours seront en langue maternelle et les sujets peuvent choisir la langue pour la réalisation des activités.
4. Formation spécifique selon la langue et la spécialité.

Chaque situation présente ses avantages et inconvénients. Peut-être que la proposition la plus réalisable consiste à créer deux matières dénommées Enseignement bilingue (anglais) et Enseignement bilingue (français) où on donne un module avec des notions générales et les futurs professeurs apprennent à élaborer des programmes en langue étrangère.

En ce qui concerne la psycholinguistique, la disponibilité pourrait aller plus loin et être étudié de manière transversale à travers de la linguistique informatique. De cette façon, il serait possible de créer un modèle informatique qui permet d'expliquer cognitivement l'ordre des mots avec de nombreuses applications pratiques. C'est le cas des épreuves de fluidité verbale qui sont amplement employées dans le cadre de la neurolinguistique et psycholinguistique et qui aide, parmi d'autres aspects, à détecter des pathologies neurologiques.

Les conclusions extraites de cette étude font un pas en avant pour des propositions et futures lignes de recherche qui contribuent au cycle de l'application

d'une méthode scientifique. Dans ce sens, le chimiste anglais Humphry Davy (1840: 362) suggère que:

For every discovery opens a new field for investigation of facts, shows us the imperfection of our theories. It has justly been said, that the greater the circle of light, the greater the boundary of darkness by which it is surrounded.

BIBLIOGRAFÍA

- Acuerdo de 22 de marzo de 2005, del Consejo de Gobierno, por el que se aprueba el Plan de Fomento del Plurilingüismo en Andalucía. BOJA núm. 65, de 5 de abril de 2005, pp. 8-39. Disponible en <http://www.juntadeandalucia.es/boja/2005/65/d5.pdf> ³²²
- Ahumada Lara, I. (2006). *El léxico disponible de los estudiantes preuniversitarios de la provincia de Jaén*. Jaén: Universidad de Jaén.
- Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. (4th Edition). West Sussex: John Wiley & Sons.
- Alaminos Chica, A. y Santacreu Fernández, O. (2010). La emigración cualificada española en Francia y Alemania. *Papers, Revista de Sociología*, 95(1), 201-211. Disponible en http://rua.ua.es/dspace/bitstream/10045/23304/1/2010_Alaminos_Santacreu_Papers.pdf
- Alba Ovalle, O. (1995a). *Léxico disponible de la República Dominicana*. Santiago de los Caballeros: Pontificia Universidad Católica Madre y Maestra.
- Alba Ovalle, O. (1995b). Anglicismos léxicos en el español dominicano: análisis cuantitativo. En O. Alba Ovalle, *El español dominicano dentro del contexto americano* (pp. 11-38). Santo Domingo: Librería La Trinitaria.
- Alba Ovalle, O. (1996). Disponibilidad léxica en el español dominicano: aspectos sociolingüísticos. En M. Arjona Iglesias (Ed.), *Actas del X Congreso Internacional de la ALFAL* (pp. 742-749). México: Universidad Nacional Autónoma de México.
- Alba Ovalle, O. (1998). Variable léxica y dialectología hispánica. *La Torre: Revista de la Universidad de Puerto Rico*, 3(7-8), 299-316.
- Alba Ovalle, O. (1999). Densidad de anglicismos en el léxico disponible de la República Dominicana. En J. A. Samper Padilla y M. Troya Déniz (Coords.), *Actas del XI Congreso Internacional de ALFAL* (pp. 853-865). Las Palmas de Gran Canaria: Publicaciones de la Universidad de Las Palmas de Gran Canaria.

³²² Todos los enlaces presentes en la bibliografía han sido consultados por última vez el 15 de marzo de 2015.

- Alba Ovalle, O. (2000a). Densidad de anglicismos en el léxico disponible. En O. Alba Ovalle, *Nuevos aspectos del español en Santo Domingo* (pp. 73-98). Santo Domingo: Librería La Trinitaria y Brigham Young University.
- Alba Ovalle, O. (2000b). Variable léxica y comparación dialectal. En O. Alba Ovalle, *Nuevos aspectos del español en Santo Domingo* (pp. 99-132). Santo Domingo: Librería La Trinitaria y Brigham Young University.
- Alba Ovalle, O. (2011). Evolución del léxico disponible de los dominicanos: observación del cambio en tiempo real. Comunicación presentada en el XVI *Congreso Internacional de la ALFAL*, 6-9 de junio, Alcalá de Henares.
- Alba Ovalle, O. (2013). Variación diacrónica del léxico disponible dominicano. *LEA: Lingüística española actual*, 35(1), 149-180.
- Albarrán Lozano, I., Guillén Estany, M. y Ayuso Gutiérrez, M. (2003). Modelo de regresión lineal múltiple y modelos de elección discreta (logit, probit, multinomiales y censurado). En J. P. Lévy Mangin y J. Varela Mallou (Coord.), *Análisis Multivariable para las Ciencias Sociales* (pp. 277-326). Madrid: Prentice Hall.
- Albert, M. L., Obler, L. K. y Obler, A. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.
- Alvar Ezquerro, M. (2000). *El Tesoro Léxico de las Hablas Andaluzas*. Madrid: Arco/Libros.
- Appel, R. y Muysken, P. (1987). *Language Contact and Bilingualism*. London: Edward Arnold.
- Arancibia Manhey, P. (2006). Indigenismos contenidos en el corpus de disponibilidad léxica de estudiantes chilenos. *Ámbitos. Revista de Estudios de Ciencias Sociales y Humanidades*, 16, 62-72.
- Areta Lara, M. (2008). *Análisis sociolingüístico de la disponibilidad léxica de los estudiantes preuniversitarios de Navarra*. (Tesis doctoral). Universidad de Navarra.
- Arnal Purroy, M.^a L. (2008). *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución Fernando el Católico.
- Arnal Purroy, M.^a L. (2008b). Niveles socioculturales y léxico dialectal en el vocabulario disponible de Aragón. En J. L. Blas Arroyo, M. Casanova Ávalos,

- M. Velando Casanova y J. Vellón Lahoz (Coord.), *Discurso y sociedad II: nuevas contribuciones al estudio de la lengua en contexto social* (pp. 457-469). Castellón de la Plana: Universitat Jaume I.
- Arnal Purroy, M.^a L. (Coord.). (2004). *Léxico disponible de Aragón*. Zaragoza: Libros Pórtico.
- Arnal Purroy, M.^a L. (Ed.). (2008a). *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución Fernando el Católico.
- Ávila Muñoz, A. M y Villena Ponsoda, J. A. (Eds.) (2010). *Variación social del léxico disponible en la ciudad de Málaga*. Málaga: Editorial Sarriá.
- Ávila Muñoz, A. M y Villena Ponsoda, J. A. (Eds.) (2011). *Estudios sobre el español de Málaga. Pronunciación, vocabulario y sintaxis*. Málaga: Editorial Sarriá.
- Ávila Muñoz, A. M. (2006). *Léxico disponible de los estudiantes preuniversitarios de Málaga*. Málaga: Universidad de Málaga.
- Ávila Muñoz, A. M. (2007). Léxico disponible y ortografía. Condicionantes sociales y hábitos culturales de influencia. En J. Moya Corral y M. Sosinski, *Las hablas andaluzas y la enseñanza de la lengua. Actas de las XII Jornadas sobre la enseñanza de la lengua española* (pp. 25-46). Granada: Universidad de Granada.
- Ávila Muñoz, A. M. y Sánchez Sáez, J. M. (2010). La disponibilidad léxica. Antecedentes y fundamentos. En A. M. Ávila Muñoz y J. A. Villena Ponsoda (Eds.), *Variación social del léxico disponible en la ciudad de Málaga* (pp. 35-81). Málaga: Editorial Sarriá.
- Ávila Muñoz, A. M. y Sánchez Sáez, J. M. (2011). La posición de los vocablos en el cálculo del índice de disponibilidad léxica: procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga. *ELUA. Estudios de Lingüística Universidad de Alicante*, 25, 45-74.
- Ávila Muñoz, A. M. y Sánchez Sáez, J. M. (2014). Fuzzy sets and Prototype Theory: Representational model of cognitive community structures based on lexical availability trials. *Review of Cognitive Linguistics*, 12(1), 133-159.
- Ayora Esteban, C. (2002). La disponibilidad léxica en una situación de contacto de lenguas. En M.^a D. Muñoz, A. I. Rodríguez-Piñero Alcalá, G. Fernández

- Smith y V. Benítez Soto (Eds.), *IV Congreso de Lingüística General, Cádiz del 3 al 6 de abril de 2000* (vol. 1, pp. 159-170). Cádiz: Universidad de Cádiz.
- Ayora Esteban, C. (2004). Deficiencias en la adquisición del léxico como problema específico de las lenguas en contacto. En M. Villayandre Llamazares (Ed.). *Actas del V Congreso de Lingüística General, León, 5-8 de marzo de 2002* (vol. 1, pp. 281-293). Madrid: Arco/Libros.
- Ayora Esteban, C. (2005). Un proyecto de léxico disponible: objetivo y utilidades. En C. Ayora Esteban y A. González Vázquez (Eds.), *Estudios de Didáctica de la Lengua y la Literatura* (pp. 49-60). Ceuta: Sección departamental de Ceuta del Departamento de Didáctica de la Lengua y la Literatura, Universidad de Granada.
- Ayora Esteban, C. (2006). *Disponibilidad léxica en Ceuta: aspectos sociolingüísticos*. Cádiz: Universidad de Cádiz.
- Ayora Esteban, C. (2007). Los estudios de disponibilidad léxica: ámbitos de aplicación. En P. Cano López (Coord.), *Actas del VI Congreso de Lingüística General, Santiago de Compostela, del 3 al 7 de mayo de 2004* (vol. 2, tomo 2, pp. 2367-2378.) Santiago de Compostela: Universidad de Santiago.
- Azurmendi Ayerbe, M.^a J. (1983). *Elaboración de un modelo para la descripción sociolingüística del bilingüismo y su aplicación parcial en la comarca de San Sebastián*. San Sebastián: Ediciones de la Caja de Ahorros Provincial de Guipúzcoa.
- Bagri, H. y Samper Padilla, J. A. (Coords.). (2013). *Disponibilidad léxica de ELE en Marruecos. Niveles de Secundaria y Enseñanza Superior en la región Souss Massa Drâa*. Agadir: Universidad Ibn Zohr.
- Baker, C. (1992). *Attitudes and Langue*. Clevedon: Multilingual Matters.
- Baker, C. (2001). *Foundations of Bilingual Education and Bilingualism* (3rd ed.). Clevedon: Multilingual Matters.
- Bakonyi, H. (1933). *Die gebräuchlichste Wörter der deutschen Sprache, für den Fremdsprachenunterricht stufenmässig zusammengestellt. Beiträge zur Methodik des deutschen Sprachunterrichts im Ausland*. Goethe-Institut zur Fortbildung ausländischer Deutschlehrer, Deutsche Akademie, Heft 1. München: Reinhardt.

- Bartol Hernández, J. A. (1998). Anglicismos en el español de finales de siglo. Anglicismos en el léxico disponible de dos regiones hispanas. En C. Kent y M. C. de la Calle (Eds.), *Visiones salmantinas (1898-1998)* (pp. 97-124). Salamanca: Universidad de Salamanca y Ohio State University.
- Bartol Hernández, J. A. (2001). Reflexiones sobre la disponibilidad léxica. En J. A. Bartol Hernández, S. Crespo Matellán, C. Fernández Juncal, C. Pensado Ruiz, E. Prieto de los Mozos y N. Sánchez González de Herrero (Eds.) *Nuevas aportaciones al estudio de la lengua española. Investigaciones filológicas* (pp. 221-236). Salamanca: Luso-Española de Ediciones.
- Bartol Hernández, J. A. (2004). *Léxico disponible de Soria. Estudio y diccionarios*. Burgos: Instituto Castellano y Leonés de la Lengua.
- Bartol Hernández, J. A. (2005). Extranjerismos disponibles: su integración y vitalidad en una comunidad de habla (provincia de Soria, España). *Spanish in Context*, 2(2), 203-229.
- Bartol Hernández, J. A. (2010). Disponibilidad léxica y selección del vocabulario. En R. M.^a Castañer Martín y V. Lagüéns Gracia (Eds.), *De moneda nunca usada: Estudios dedicados a José M.^a Enguita Utrilla* (pp. 85-107). Zaragoza: Institución Fernando el Católico. Disponible en <http://ifc.dpz.es/recursos/publicaciones/29/95/09bartol.pdf>
- Becker, H. S., Geer, B., Hughes, E. C. y Strauss, A. L. (1961). *Boys in white: Student culture in medical school*. Chicago: University of Chicago Press.
- Bellón Fernández, J. J. (2011). *Léxico disponible de la provincia de Córdoba*. Córdoba: Servicio de Publicaciones de la Universidad de Córdoba.
- Benítez Pérez, P. (1992). Disponibilidad léxica en la zona metropolitana de Madrid. *Boletín de la Academia Puertorriqueña de la Lengua Española*, 1(1), 71-102.
- Benítez Pérez, P. (1994a). Léxico real/irreal en los manuales de español para Extranjeros. En S. Montesa Peydró y A. Garrido Moraga (Eds.), *Actas del II Congreso Nacional de ASELE. Español para extranjeros: didáctica e investigación* (pp. 325-333). Málaga: ASELE.
- Benítez Pérez, P. (1994b). Convergencias y divergencias en el léxico de alumnos de COU. *Revista de Estudios de Adquisición de la Lengua Española*, 2, 39-45.

- Benítez Pérez, P. (1997). El vocabulario enseñado en los manuales de ELE, ¿es el adecuado? En P. Jiménez, F. Vanoverberghe y R. Verdonk (Eds.), *Problemas actuales en la enseñanza del español como lengua extranjera: gramática, pragmática, vocabulario y cultura* (pp. 64-76). Leuven: Wolters.
- Benítez Pérez, P. y Zebrowski, J. (1993). El léxico español en los manuales polacos. En S. Montesa Peydró y Garrido Moraga, A. (Eds.), *Actas del III Congreso Nacional de la Asociación para la enseñanza del español como lengua extranjera (ASELE). El español como lengua extranjera: de la teoría al aula* (pp. 223-230). Málaga: ASELE.
- Bernstein, B. (1961). Social structure, language and learning. *Educational Research* 3, 163-173.
- Bernstein, B. (1971-1975). *Class, Codes and Control* (I-III). London: Routledge and Kegan Paul.
- Bialystok, E. (Ed.). (1991). *Language processing in bilingual children*. Cambridge: Cambridge University Press.
- Birdsong, D. (2005). Interpreting age effects in second language acquisition. En J. F. Kroll y A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 109-127). New York: Oxford University Press.
- Blanco Domínguez, M. (2011). La ortografía en el léxico disponible del español de Galicia. En B. López Meirama (Ed.), *Estudios sobre disponibilidad léxica en el español de Galicia* (pp. 189-216). Santiago de Compostela: Universidad de Santiago de Compostela.
- Blasco Mateo, E., Torres Torres, A., Forment Fernández, M.^a M. y Pérez Edo, M.^a Á. (2003). La disponibilidad léxica en la provincia de Barcelona: aspectos metodológicos [en prensa]. *XXXIII Simposio de la Sociedad Española de Lingüística*. Gerona, 16-19 de diciembre de 2003.
- Bombarelli, Á. (2005). *La disponibilidad léxica como herramienta didáctica: una propuesta de selección del vocabulario para un nivel umbral de ELE*. (Memoria de Máster inédita). Universidad de Salamanca.
- Bongers, H. (1947). *The History and principles of vocabulary control as it affects the teaching of foreign languages in general and of English in particular*. Woerden, Holland: Wocopi.

- Bonilla Rengifo, L. (2004). *Disponibilidad léxica en Popayán* (Colombia). (Tesina inédita). Universidad de Salamanca.
- Borrego Nieto, J. (2004). Algunas preguntas en relación con el concepto de léxico disponible. En J. Prado Aragonés y M.^a V. Galloso Camacho (Coords.), *Diccionario, léxico y cultura* (pp. 59-69). Huelva: Universidad de Huelva.
- Borrego Nieto, J. y Fernández Juncal, C. (2002). Léxico disponible: aplicaciones a los estudios dialectales. En M.^a D. Muñoz, A. I. Rodríguez-Piñero Alcalá, G. Fernández Smith y V. Benítez Soto (Eds.), *IV Congreso de Lingüística General* (vol. 2, pp. 297-306), Cádiz del 3 al 6 de abril de 2000. Cádiz: Universidad de Cádiz.
- Borrego Nieto, J. y Fernández Juncal, C. (2003). ¿En qué cambia la universidad la disponibilidad léxica de los preuniversitarios? En F. Moreno Fernández, J. A. Samper Padilla, M. Vaquero, M.^a L. Gutiérrez Araus y F. Gimeno Menéndez (Coords.), *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales* (vol. 1, pp. 167-178). Madrid: Arco/Libros.
- Braun, A. (2007). Immersion et Compréhension en Lecture. En L. Puren y S. Babault (Eds.), *L'Éducation au-delà des Frontières* (pp. 215-257). Paris: L'Harmattan.
- Brecht, R. D. y Robinson, J. L. (1995). On the Value of Formal Instruction in Study Abroad: Student Reactions in Context. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 317-333). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Brecht, R., Davidson, D. y Ginsberg, R. (1995). Predictors of Foreign Language Gain During Study Abroad. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*, pp. 37-66. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Brewster, J., Ellis, G. y Girard, D. (1992). *The Primary English Teacher's Guide*. Harmondsworth: Penguin.
- Brown, R., Waring, R., y Donkaewbua, S. (2008) Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136-163.

- Butrón, G. (1987). *El léxico disponible: índices de disponibilidad*. (Tesis doctoral inédita). Universidad de Puerto Rico.
- Butrón, G. (1989). Aspectos sociolingüísticos de la disponibilidad léxica. *Asomante*, 1-2, 29-37.
- Butrón, G. (1991). Nuevos índices de disponibilidad léxica. En H. López Morales (Ed.), *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre «Aportes de la lingüística a la enseñanza del español como lengua materna»* (pp. 79-89). Río Piedras: Universidad de Puerto Rico.
- Calvo Gómez, F. (1993). *Técnicas estadísticas multivariantes: con resolución de ejercicios prácticos mediante los paquetes estadísticos SPSS y PROGSTAD*. País Vasco: Universidad de Deusto.
- Camacho Rosales, J. (2002). *Estadística con SPSS para Windows versión 11*. Madrid: Rama.
- Cañizal Arévalo, A. (1987): *Disponibilidad léxica en escolares de primaria terminada. Análisis de seis centros de interés*. (Memoria de licenciatura inédita). Universidad Autónoma de México.
- Cañizal Arévalo, A. (1991). Redes semánticas y disponibilidad léxica en el español de escolares mexicanos. En C. Hernández Alonso, G. P. Granda, C. Hoyos, V. Fernández, D. Dietrick e Y. Carballera (Eds.), *El español de América. Actas del III Congreso Internacional de 'El español de América'* (vol. II, pp. 631-637). Valladolid: Junta de Castilla y León.
- Caramazza, A. (1999). Neural Basis of Lexicon, En R. Wilson y F. Kiel (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*, 469-471. Cambridge, MA: The MIT Press.
- Disponible en <http://www.ru.lv/~peter/zinatne/ebooks/MIT.Encyclopedia.of.the.Cognitive.Sciences.pdf>
- Carcedo González, A. (1998a). Sobre las pruebas de disponibilidad léxica para estudiantes de español/LE. En M.^a V. Romero Gualda (Ed.), *Español como lengua extranjera: investigación y docencia*, RILCE: Revista de Filología Hispánica, 14(2), 205-224.

- Carcedo González, A. (1998b). Tradición y novedad en las aportaciones hispánicas a los estudios de disponibilidad léxica, *Lingüística*, 10, 5-68.
- Carcedo González, A. (1999a). Vocabularios básicos, fundamentales y disponibles. Renovación metodológica de la investigación hispánica. En Y. Gambier y E. Suomela-Salmi (Eds.), *Jalons pour le 75e anniversaire de l'enseignement du français à l'Université de Turku. Publications du Département d'Études Françaises*, 2. (pp. 191-207). Turku: Université de Turku.
- Carcedo González, A. (1999b). Estudio comparativo del vocabulario español (LE) disponible de estudiantes finlandeses y el de la sintopía madrileña: propuestas didácticas. *Documentos de Español Actual*, 1, 73-87.
- Carcedo González, A. (1999c). Análisis de errores léxicos del español en la interlingua de los finlandeses. En M.^a C. Losada Aldrey, J. F. Márquez Caneda y T. E. Jiménez Juliá (Coords.), *Español como lengua extranjera: enfoque comunicativo y gramática. Actas del IX Congreso Internacional de ASELE* (pp. 465-472). Santiago de Compostela: Universidad de Santiago de Compostela.
- Carcedo González, A. (2000a). Índices léxico-estadísticos y graduación del vocabulario en la enseñanza de ELE (aspectos culturales). En M. Franco Figueroa Franco, C. Soler, J. De Cos, M. Rivas y F. Ruiz (Eds.), *Nuevas perspectivas en la enseñanza del español como lengua extranjera Actas del X Congreso Internacional de ASELE* (vol. 1, pp. 174-183). Cádiz: Universidad de Cádiz.
- Carcedo González, A. (2000b). La lengua como manifestación de otredad cultural (o convergencia intercultural [versión electrónica]. En J. M. Aguirre Romero (Ed.), *Especulo. Cultura e Intercultural en la enseñanza del español como lengua extranjera*. Madrid: Universidad Complutense de Madrid. Disponible en <http://www.ucm.es/info/especulo/ele/carcedo.html>
- Carcedo González, A. (2000c). *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Turku: Turun Yliopisto.
- Carcedo González, A. (2001). *Léxico disponible de Asturias*. Turku: Universidad de Turku.

- Carcedo González, A. (2003). Unidad y variedad diatópica de la disponibilidad léxica del español: comparación de los inventarios de Puerto Rico, Cádiz y Asturias. En F. Moreno Fernández, J. A. Samper Padilla, M. Vaquero, M.^a L. Gutiérrez Araus, y F. Gimeno Menéndez (Coords.), *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales* (vol. 1, pp. 199-225). Madrid: Arco/Libros.
- Carcedo González, A. (2004). La variable léxica disponible en la comparación interdialectal: compatibilidad de la norma asturiana con otras sintopías hispanohablantes. En V. Sánchez Corrales (Ed.), *Actas del XIII Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina* (pp. 661-679). San José de Costa Rica: Universidad de Costa Rica.
- Carlson, J., Burn, B., Ussem, J. y Yachimowicz, D. (1990). *Study abroad. The experience of American undergraduates in Eastern Europe and the United States*. West port, Conn.: Greenwood Press.
- Carrera de la Red, M.^a F. y Bradley, W. E. (2004a). El proyecto sobre disponibilidad léxica en el mundo hispánico. Su aplicación en Cantabria. En G. Martínez Camino, M. García Casado y S. Scurfield (Eds.), *Encuentro de lenguas y literaturas* (pp. 21-30). Santander: Universidad de Cantabria.
- Carrera de la Red, M.^a F. y Bradley, W. E. (2004b). Extranjerismos en el léxico disponible de Cantabria. En M. Villayandre Llamazares (Ed.), *Actas del V Congreso de Lingüística General* (vol. 1, pp. 535-545). Madrid: Arco/Libros.
- Carroll, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1(2), 131-151.
- Casanova Ávalos, M. (2006). La disponibilidad léxica en la Comunidad Valenciana. En J. L. Blas Arroyo, M. Casanova Ávalos y M. Velando Casanova (Eds.), *Discurso y sociedad: contribuciones al estudio de la lengua en contexto social* (pp. 737-751). Castellón de la Plana: Universidad Jaume I.
- Casanova Ávalos, M. (2008). Anglicismos en el léxico disponible de Castellón. En J. L. Blas Arroyo, M. Casanova Ávalos, M. Velando Casanova y J. Vellón (Eds.), *Discurso y sociedad II: nuevas contribuciones al estudio de la lengua en context social* (pp. 471-499). Castellón de la Plana: Universidad Jaume I.

- Castañer Martín, R. M.^a (2008). Los extranjerismos en el léxico disponible de los jóvenes aragoneses. En M.^a L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 51-74). Zaragoza: Institución Fernando el Católico.
- Castro Calvín, J. (1991). A Reappraisal of Motivation: The Cornerstone of L2 Language Learning. *Revista Alicantina de Estudios Ingleses*, 4, 11-23.
- Cenoz Iragui, J. (1993). Diferencias individuales en la adquisición del inglés. *Revista Española de Lingüística Aplicada*, 9, 27-36.
- Cenoz, J. (2000). Research on Multilingual Acquisition. En J. Cenoz y U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (Vol. 19, pp. 39-53). Clevedon: Multilingual Matters.
- Centre nationale de la recherche scientifique y Université de Lorraine. *Trésor de la Langue Française* [en línea]. Disponible en <http://atilf.atilf.fr>
- Cerdà Massó, R. (Coord.). (1986). *Diccionario de lingüística*. Madrid: Ediciones Generales Anaya.
- Clyne, M. (1997). Some of the things trilinguals do. *International Journal of Bilingualism*, 1, 95-116.
- Coady, J. 2000. L2 vocabulary acquisition through extensive reading. En J. Coady y T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 225-237). Cambridge: Cambridge University Press.
- Coffey, A. y Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks, CA: Sage.
- Cohen, M. (1955). *Français élémentaire? Non*. Paris: Éditions Sociales.
- Coleman, J. y Rouxville, A. (Eds.). (1993). *Integrating new approaches*. London: Association for French Studies in association with the Center for Information on Language Teaching and Research.
- Collentine, J. (2009). Study abroad research: Findings, implications and future directions. En C. Doughty y M. Long (Eds.), *Language learners in study abroad contexts* (pp. 120-150). Clevedon: Multilingual Matters.
- Collins Concise English Dictionary*. (2008). (7th Edition). Glasgow: HarperCollins Publishers Ltd.

- Collins Spanish Dictionary. Complete and Unabridged.* (2005). (8th Edition). Glasgow: HarperCollins Publishers Ltd.
- Collins, A. y Quillian, M. (1972). How to Make a Language User. En E. Tulving y W. Donaldson (Eds.), *Organization and Memory* (pp. 309-351). New York: Academic Press.
- Colombo, N. V., Trevisán, O. N. y Wingeyer, H. R. (2013). Aportes para las representaciones sobre la sexualidad en la disponibilidad léxica de estudiantes secundarios de la región NEA. *De Prácticas y Discursos. Cuadernos de Ciencias Sociales*, 2, 1-17. Disponible en <http://biblioteca.clacso.edu.ar/gsd/collect/ar/ar-010/index/assoc/D10147.dir/ColomboTrevisanWingeyer.pdf>
- Commission Européenne. (1995). *Livre blanc sur l'éducation et la formation. Enseigner et apprendre. Vers la société cognitive.* Disponible en http://europa.eu/documents/comm/white_papers/pdf/com95_590_fr.pdf
- Comte, A. (1896). *The positive philosophy of Auguste Comte*. Traducido y editado por Harriet Martineau. London: George Bell and Sons.
- Consejo Asesor para la Segunda Modernización de Andalucía (2003). *Estrategias y propuestas para la segunda modernización de Andalucía*. Sevilla: Junta de Andalucía. Disponible en <http://www.redes-cepalcala.org/inspector/DOCUMENTOS%20Y%20LIBROS/TIC/librosegundamodernizacion.pdf>
- Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Traducido por el Instituto Cervantes. Madrid: Ministerio de Educación Cultura y Deporte y Anaya. Disponible en http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/
- Corominas, J. (1954). *Diccionario crítico etimológico de la lengua castellana*. Madrid: Editorial Gredos. Madrid.
- Cruzado Romero, J. (2003). *El léxico disponible de alumnos de 2.º, 4.º de ESO y 2.º de Bachillerato-LOGSE de Moguer (Huelva)*. (Tesina inédita). Universidad de Huelva.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of educational research*, 49(2), 222-251.

- Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. *Language processing in bilingual children*, 70-89.
- Davie, J. (1996). Language skills, course development and the year abroad. *Language Learning Journal*, 13, 73-76.
- Davy, H. (1840). *Salmonia: or, Days of fly-fishing; in a series of conversations: with some account of the habits of fishes belonging to the genus Salmo. Consolation in travel, or, The last days of a philosopher*. London: Smith, Elder & Co. Disponible en <http://babel.hathitrust.org/cgi/pt?id=hvd.32044023814494;view=1up;seq=10>
- De Gortari, E. (1979). *El método de las ciencias sociales. Nociones elementales*. México: Grijalbo.
- Decreto 327/2010, de 13 de julio, por el que se aprueba el Reglamento Orgánico de los Institutos de Educación Secundaria. BOJA núm. 139 de 16 de julio de 2010, pp. 8-34. Disponible en <http://www.juntadeandalucia.es/boja/2010/139/d2.pdf>
- Decreto 329/2010, de 13 de julio, por el que se autoriza la implantación de enseñanzas universitarias de Grado, Máster y Doctorado, se actualiza el Catálogo de Enseñanzas Universitarias conducentes a la expedición por las Universidades Públicas de Andalucía de títulos oficiales, así como la estructura de los centros que las imparten y se fijan los precios públicos y tasas a satisfacer por la prestación de servicios académicos y administrativos universitarios para el curso 2010/2011. BOJA núm. 139 de 16 de julio de 2010, pp. 57-97. Disponible en <http://www.juntadeandalucia.es/boja/2010/139/d4.pdf>
- DeKeyser, R. M. (1991). Foreign language development during a semester abroad. En B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 104-119). Lexington, MA: D. C. Heath.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533.
- Del Pino, E. (2006). *Diccionario del habla malagueña* (2.^a ed.). Córdoba: Editorial Almuzara.

- Deutscher, I. (1973). *What We Say/What We do: Sentiments and Acts*. Glenview, IL: Scott Foresman.
- Deutscher, I. (1973). *What we say/what we do: Sentiments and acts*. Glenview, Illinois: Scott, Foresman
- Dewey, D. P. (2008). Japanese vocabulary acquisition by learners in three contexts. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 15, 127-148.
- Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26, 249-273.
- Díaz-Campos, M., y Lazar, N. (2003). Acoustic analysis of voiceless initial stops in the speech of Study Abroad and Regular Class students: Context of learning as a variable in Spanish second language acquisition. En P. Kempchinsky y C. E. Piñeros, (Eds.), *Theory, practice, and acquisition: Papers from the 6th Hispanic Linguistics Symposium and the 5th Conference on the Acquisition of Spanish and Portuguese* (pp. 352-370). Somerville: Cascadilla Press.
- Diccionario Espasa Grand: español-francés français-espagnol*. (2000). Madrid: Espasa-Calpe S.A.
- Dimitrijević, N. (1969). *Lexical Availability. A new aspect of the lexical availability of secondary school children*. Heidelberg: Julius Gross Verlag.
- Dirección de Programación Académica Universidad Nacional del Litoral, Secretaría Académica. (2008). *Autoevaluación institucional de la UNL. Segunda Fase*. Santa Fe, Argentina: Centro de Publicaciones UNL.
- Doidge, N. (2007). *The Brain that changes itself: stories of personal triumph from the frontiers of brain science*. New York: Penguin Books.
- Donalson, M. (1987). *Children's minds*. Buckingham: The Open University Press.
- Dörnyei, Z. (1994a). Motivation and Motivating in the Foreign Language Classroom. *The Modern Language Journal*, 78(3), 273-283. Disponible en <http://www.zoltandornyei.co.uk/uploads/1994-dornyei-mlj-a.pdf>
- Dörnyei, Z. (1994b). Understanding L2 motivation: On with the challenge! *The Modern Language Journal*, 78(4), 515-523. Disponible en <http://www.zoltandornyei.co.uk/uploads/1994-dornyei-mlj-b.pdf>

- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Dörnyei, Z. y Csizér, K. (2012). How to design and analyze surveys in SLA research? En A. Mackey y S. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 74-94). Malden, MA: Wiley-Blackwell. Disponible en <http://www.zoltandornyei.co.uk/uploads/2012-dornyei-csizer-rmsla.pdf>
- Dörnyei, Z. y Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing* (2nd ed.). New York: Routledge.
- Dubois, J., Giacomo, M., Guespin, L. y Marcellesi, C. (1973). *Dictionnaire de Linguistique*. Paris: Larousse.
- Dupré, J. B. y Claudel, L. (2007). Francophonies minoritaires et disponibilité lexicale: le cas de jeunes locuteurs de Saint-Lazare (Manitoba) et d'Ottawa (Ontario). *Cahiers franco-canadiens de l'Ouest*, 19(2), 159-176.
- Durkheim, É. (1938). *The Rules of Sociological Method*. Chicago: University of Chicago Press.
- Durkheim, É. (1951). *Suicide. A Study in Sociology* (J. A. Spaulding y G. Simpson, trans.). Glencoe, IL: The Free Press of Glencoe. (Original work published 1897).
- Duverger, J. y Maillard, J. P. (1996). *L'enseignement bilingue aujourd'hui*. Saint-Amand-Montrond: Albin Michel.
- Echeverría, M. S. (1991). Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio. En H. López Morales (Ed.), *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre «Aportes de la lingüística a la enseñanza del español como lengua materna»* (pp. 61-78). Río Piedras: Universidad de Puerto Rico.
- Echeverría, M. S. (2000). Desarrollo de la metacognición lingüística mediante apoyo computacional. *RLA. Revista de lingüística teórica y aplicada*, 38, 61-74.
- Echeverría, M. S. (2001a). Estructura y funciones de un software de vocabulario disponible. *RLA. Revista de lingüística teórica y aplicada*, 39, 87-100.

- Echeverría, M. S. (2001b). Variación dialectal del español: software multimedial para su conocimiento. *Estudios Filológicos*, 36, 117-127. Disponible en <http://www.redalyc.org/articulo.oa?id=173413831008>
- Echeverría, M. S. (2002). Programas computacionales para el español como lengua materna. *Revista signos*, 35(51-52), 163-193.
- Echeverría, M. S. y Ramos, M. (2002). Antex '98: Un tutorial interactivo para el análisis de textos. En G. Parodi (Ed.), *Lingüística e interdisciplinariedad: Desafíos del nuevo milenio. Ensayos en honor a Marianne Peronard* (pp. 375-385.). Valparaíso, Chile: Ediciones Universitarias de Valparaíso UCV.
- Echeverría, M. S. y Vargas Opitz, R. (2006). *DispoGrafo. Programa computacional para el análisis de relaciones semánticas*. Concepción, Chile: Universidad de Concepción (Software).
- Echeverría, M. S., Herrera, M.^a O., Moreno, P. y Pradenas, F. (1987). Disponibilidad léxica en Educación Media. *RLA. Revista de Lingüística Teórica y Aplicada*, 25, 55-115.
- Echeverría, M. S., Urzúa, P y Figueroa, I. (2005). *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica*. Concepción: Universidad de Concepción.
- Echeverría, M. S., Urzúa, P. y Ferreira, R. (2008). DispoGrafo. Una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible. *RLA. Revista de lingüística teórica y aplicada*, 46, 81-91.
- El Colegio de México, A. C. (2010). *Diccionario del Español de México (DEM)*. México: El Colegio de México, El Colegio de México, Centro de Estudios Lingüísticos y Literarios.
- Elley, W. y Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19, 53-67.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Enguita Utrilla, J. M.^a (2008). Norma urbana y norma rural en el léxico disponible de los jóvenes aragoneses. En M.^a L. Arnal Purroy (Ed.), *Estudios sobre*

- disponibilidad léxica en los jóvenes aragoneses* (pp. 75-102). Zaragoza: Institución Fernando el Católico.
- Escoriza Morera, L. (2003). Lenguas en contacto y disponibilidad léxica: la situación lingüística e intercultural de Ceuta y Gibraltar. *Linred: Revista electrónica de lingüística*, 1 [en línea]. Disponible en http://www.linred.es/informacion_pdf/informacion1_271003.pdf
- Escoriza Morera, L. (2007). Disponibilidad léxica y multilingüismo. El contacto entre inglés y español en Gibraltar. En P. Cano López, I. Fernández López, M. González Pereira, G. Prego Vázquez y M. Souto Gómez, M. (Eds.), *Actas del VI Congreso de Lingüística General, Santiago de Compostela, 3-7 de mayo de 2004* (vol. 2, t. 2, Las lenguas y su estructura, pp. 2477-2484). Madrid: Arco/Libros.
- Etxebarria Arostegui, M. (1996). Disponibilidad léxica en escolares del País Vasco. Variación sociolingüística y modelos de enseñanza bilingüe. *Revista española de lingüística*, 26(2), 301-325. Disponible en <http://www.sel.edu.es/pdf/jul-dic-96/26-2-Etxebarria.pdf>
- Etxebarria Arostegui, M. (1999). Disponibilidad léxica y enseñanza de la lengua materna en el área metropolitana del Gran Bilbao. En J. A. Samper Padilla y M. Troya Déniz (Coords.), *Actas del XI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina* (Vol. 2, pp. 1479-1494). Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria.
- Etxeberria, J., Joaristi, L. y Lizasoain, L. (1990). *Programación y análisis estadísticos básicos con SPSS/PC +*. Madrid: Paraninfo.
- Eustache F. (1993). Langage, vieillissement et démences. En F. Eustache, *Langage et aphasie*, 205-227. Bruxelles: De Boeck.
- Fall, K. (1976). *L'univers Culturel de l'enfant wolophone de Dakar, d'après la disponibilité des concepts dans la langue maternelle et la langue d'enseignement (cas du wolof et du français)*. (Thèse de maîtrise). Université Laval, Québec.
- Fasce, E., Echeverría, M. S., Matus, O., Ortiz, L., Palacios, S. y Soto, A. (2009). Atributos del profesionalismo estimados por estudiantes de medicina y médicos. Análisis mediante el modelo de disponibilidad léxica. *Revista Médica de*

Chile, 137, 746-752. Disponible en

<http://www.scielo.cl/pdf/rmc/v137n6/art03.pdf>

Fernández Fernández, S., Cordero Sánchez, J. M.^a y Córdoba Largo, A. (2002).

Estadística descriptiva (2.^a ed.). Madrid: ESIC Editorial.

Fernández Juncal, C. (2008). *Léxico disponible de Burgos*. Burgos: Fundación Instituto

Castellano y Leonés de la Lengua.

Fernández Juncal, C. (2013). *Léxico disponible de Cantabria*. Salamanca: Ediciones

Universidad de Salamanca.

Fernández Smith, G., Rico Martín, A. M.^a, Molina García, M.^a J. y Jiménez Jiménez,

M.^a Á. (2008). *Léxico disponible de Melilla: estudio sociolingüístico y repertorios léxicos*.

Madrid: Arco/Libros.

Fernández Smith, G., Rico Martín, A. M.^a, Molina García, M.^a J. y Jiménez Jiménez,

M.^a Á. (2006). Proyecto sobre disponibilidad léxica en alumnos

preuniversitarios de Melilla. En J. L. Blas Arroyo, M. Casanova Ávalos, y M.

Velando Casanova (Eds.), *Discurso y sociedad: contribuciones al estudio de la lengua en contexto social* (pp. 767-777). Castellón de la Plana: Universidad Jaume I.

Fernández-Merino Gutiérrez, P. V. (2015). *Léxico disponible de inmigrantes en Castilla y*

León. Estudio y diccionarios. Asturias: SAPERE AUDE.

Ferreira, R. A. (2006). *Disponibilidad léxica en inglés como lengua materna e inglés como lengua*

extranjera. Estudio del léxico disponible desde un enfoque psicolingüístico. (Tesis de magister inédita). Universidad de Concepción.

Ferreira, R., y Echeverría, M. S. (2010-2011). Redes semánticas en el léxico disponible

de inglés L1 e inglés LE. *Onomázein*, 21, 133-153.

Finegan, E. (1994). *Language. Its Structure and Use*. Forth Worth: Harcourt Brace.

Flick, U. (2002). *An Introduction to Qualitative Research* (2nd ed.). London: Sage.

Fodor, J. D., Fodor, J. A. y Garrett, M. (1975). The psychological unreality of

semantic representations. *Unguistic Inquiry*, 4, 515-31.

Foster, P. (2009). Lexical diversity and native-like selection: The bonus of studying

abroad. En H. Daller, D. Malvern, P. Meara, J. Miltron, B. Richards y J.

Treffers-Daller, *Vocabulary studies in first and second language acquisition: The*

- interface between theory and application* (pp. 91-106). Basingstoke: Palgrave MacmillanB.
- Fox W. y Bayat, M. S. (2007). *A guide to managing research*. Ciudad del Cabo: Juta and Company Ltd.
- Freed, B. (1995b). What makes us think that students who study abroad become fluent? En B. Freed (Ed.), *Second language acquisition in a study abroad context*, pp. 123-148. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Freed, B. (Ed.). (1995a). *Second language acquisition in a study abroad context*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Freed, B., Segalowitz, N. y Dewey, D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(2), 275-301.
- Frey Pereyra, M.^a L. H. (2006). Disponibilidad léxica y escritura del ELE: propuesta de comparación de dos corpus. *Interlingüística*, 21, 366–373.
- Frey Pereyra, M.^a L. H. (2008). *El léxico disponible en los escritos de alumnos de español como lengua extranjera. Estudio comparativo de los corpóra*. (Tesis doctoral). Universidad de Alcalá.
- Fuentes Páez, C. (2014). Identificando y trabajando con el léxico disponible. Experiencias iniciales [en línea]. Comunicación presentada en el *V Simposio para el mejoramiento de la enseñanza y el aprendizaje de las lenguas, una necesidad formativa urgente de atender para el logro de la educación omnilateral en el siglo XXI*, 12-14 de mayo, Universidad Nacional Autónoma de México, México, D. F. Consultado en http://cel.prepa2.unam.mx/documentos/eventos/5osimposio/M2_14.pdf
- Fuster Pérez, Nadja N. (2013). *El léxico disponible de jóvenes puertorriqueños: Reflejo sociocultural de dos épocas*. (Tesina de máster en Lingüística y Lexicografía inédita). Universidad de León, Escuela de Lexicografía Hispánica, Asociación de Academias de la Lengua Española.
- Gajo, L. y Serra, C. (2002). Bilingual Teaching: Connecting Language and Concepts in Mathematics. En W. C. Daniel y M. G. Jones (Eds.), *Education and Society in*

- Plurilingual Contexts* (pp. 75-95). Brussels: VUB Brussels University Press.
- Galisson, R. (1971). *Inventaire thématique et syntagmatique du français fondamental*. Paris: Hachette.
- Galisson, R. (1979). *Lexicologie et enseignement des langues*. Paris: Hachette.
- Galisson, R. (1991). *De la langue à la culture par les mots*. Paris: Clé Internationale.
- Gallardo del Puerto, F., y Martínez Adrián, M. (2014). The Incidence of Previous Foreign Language Contact in a Lexical Availability Task: A Study of Senior Learners. En R. Jiménez Catalán, *Lexical Availability in English and Spanish as a Second Language* (pp. 53-68). Dordrech: Springer Netherlands.
- Gallego Gallego, D. J. (2011). Disponibilidad léxica en estudiantes de español como lengua extranjera en la Comunidad de Madrid. Aspectos metodológicos y primeros resultados [edición en CD]. En A. M.^a Cestero Mancera, I. Molina Martos y F. Paredes García (Eds.) *Documentos para el XVI Congreso Internacional de la ALFAL*. Alcalá de Henares: Universidad de Alcalá.
- Gallego Gallego, D. J. (2014). Léxico disponible de estudiantes de español como lengua extranjera en la comunidad de Madrid. (Tesis doctoral inédita). Universidad de Alcalá.
- Galloso Camacho, M.^a V. (1998). *El léxico disponible en el nivel preuniversitario (provincia de Zamora)*. (Memoria de licenciatura). Universidad de Salamanca.
- Galloso Camacho, M.^a V. (2002a). *El léxico de los estudiantes preuniversitarios en el distrito universitario de Salamanca*. Salamanca: Universidad de Salamanca.
- Galloso Camacho, M.^a V. (2002b). *El léxico disponible en el nivel preuniversitario. Provincia de Zamora*. Huelva: Universidad de Huelva.
- Galloso Camacho, M.^a V. (2003). *El léxico disponible de Ávila, Salamanca y Zamora*. Burgos: Instituto Castellano y Leonés de la Lengua.
- García Casero, M. (2013). *El léxico disponible en estudiantes de 4.º de Educación Secundaria Obligatoria en Santander*. (Tesis doctoral inédita). Universidad de Cantabria.
- Disponible en <https://www.educacion.gob.es/tesco/imprimirFicheroTesis.do?fichero=42629>

- García Domínguez, M. J., Marrero Pulido, V., Pérez Martín, J. A. y Piñero Piñero, G. (1994). Estudio de la disponibilidad léxica de Gran Canaria. La variable geográfica y el tipo de educación. *Revista Estudio Adquisición de la Lengua Española (REALE)*, 2, 65-72.
- García Laborda, J. y Bejarano, L. G. (2008). Adquisición de una segunda lengua en estancias cortas en el extranjero: un análisis actitudinal. *Didáctica. Lengua y literatura*, 20, 117-134.
- García Marcos, F. J. y Mateo García, M.^a V. (2000). *La selección de materiales léxicos en la enseñanza de lenguas extranjeras. Evaluación y propuestas desde la disponibilidad léxica*. MS.
- García Megía, A. (2003). *La disponibilidad léxica en la ciudad de Almería entre los nueve y los doce años*. (Tesis doctoral inédita). Universidad de Almería.
- García Roldán, J. (1995). *Cómo elaborar un proyecto de investigación*. Alicante: Secretariado de Publicaciones, Universidad de Alicante.
- García Santa Cecilia, Á. (1995). *El currículo de español como lengua extranjera*. Madrid: Edelsa.
- García Santa Cecilia, Á. (2000). *Cómo se diseña un curso de lengua extranjera*. Madrid: Arco/Libros.
- García Yebra, V. (1982). *Teoría y práctica de la traducción*. Madrid: Gredos.
- Gardner, R. C. (1982). Language attitudes and language learning. En H. Giles y E. Bouchard Ryan (Eds.), *Attitudes towards language variation* (pp. 132-147). London: Edward Arnold.
- Gardner, R. C. y Lambert, W. E. (1972): *Attitudes and Motivation in Second Language Learning*. Rowley, MA: Newbury House.
- Garnham, A. (1992). *Psycholinguistics: Central Topics*. London-Nueva York: Routledge.
- Gass, S. y Selinker, L. (2001). *Second Language Acquisition: An introductory course* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Germany, P. G. y Cartes, N. E. 2000. Léxico disponible en inglés como segunda lengua en instrucción formalizada, en *Estudios Pedagógicos (Valdivia)*, 26, 39-50. Disponible en**

http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07052000000100003

- Gómez Devís, M.^a B. (2004). *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: reflexión metodológica, análisis sociolingüístico y aplicaciones*. Valencia: Universidad de Valencia. Disponible en: <http://www.tdx.cat/bitstream/handle/10803/9814/gomez.pdf?sequence=1>
- Gómez Devís, M.^a B. y Llopis Rodrigo, F. (2010b). Léxico disponible del catalán de Valencia: planificación lingüística de una lengua minoritaria. *Romanitas, lenguas y literaturas romances*, 4(2) [en línea]. Disponible en <http://romanitas.uprrp.edu/espanol/volumen4/DEF.pdf>
- Gómez Molina, J. R. (2006). La variable 'sexo' en los estudios de disponibilidad léxica. *RLA, Revista de Lingüística Teórica y Aplicada*, 44: 47-65.
- Gómez Molina, J. R. y Gómez Devís, B. (2004). *La disponibilidad léxica de los estudiantes preuniversitarios valencianos. Estudio de estratificación sociolingüística*. Valencia: Universitat de València.
- Gómez Sánchez, M.^a E. y Guerra Salas, L. (2005a). El español de los profesionales de la radio. Una aproximación desde la disponibilidad léxica. En M.^a L. Carrión (Ed.), *Perspectivas interdisciplinares de la Lingüística Aplicada. Actas del XXII Congreso de AESLA* (pp. 351-360). Valencia: Universidad Politécnica.
- Gómez Sánchez, M.^a E. y Guerra Salas, L. (2005b). Disponibilidad léxica y fines específicos: análisis del centro de interés 'prensa'. En I. Sanz Sainz y Á. Felices Lagos (Eds.), *Las nuevas tendencias de las lenguas de especialidad en un contexto internacional y multicultural* (pp. 695-702). Granada: Universidad de Granada.
- Gontier, D. (1986). *Exécution, traitement, analyse et commentaire d'une enquête de disponibilité lexicale en milieu scolaire senegalais (classes de CM2, 5^{ème} et 3^{ème})*. (Thèse de troisième cycle). Université de Strasbourg II.
- González Fernández, J. (2013). La disponibilidad léxica de los estudiantes turcos de español como lengua extranjera [en línea]. *MarcoELE*, 16. Disponible en <http://marcoele.com/disponibilidad-lexica-de-estudiantes-turcos/>
- González Fernández, J. (2014). Idoneidad de los centros de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera

- [en línea]. *Revista Nebrija de Lingüística Aplicada*, 16. Disponible en http://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_52f65e512df78.pdf
- González Martínez, A. (2000-2001). Compatibilidad y discrepancia entre los léxicos disponibles de Puerto Rico y Cádiz. *Pragmalingüística*, 8-9, 179-196.
- González Martínez, A. (2002). *La disponibilidad léxica de los alumnos preuniversitarios de la provincia de Cádiz*. Cádiz: Universidad de Cádiz.
- Goodwin, C. y Nacht, M. (1988). *Abroad and beyond: Patterns in American overseas expansion*. Cambridge: Cambridge University Press.
- Gougenheim, G. (1956). Français élémentaire, oui ou non? *Bulletin de la Faculté des Lettres de Strasbourg*, 8-9.
- Gougenheim, G., Michéa, R., Rivenc, P. y Sauvageot, A. (1956). *L'élaboration du français élémentaire*. Paris: Didier.
- Gougenheim, G., Michéa, R., Rivenc, P. y Sauvageot, A. (1964). *L'élaboration du français fondamental (I degré). Étude sur l'élaboration d'un vocabulaire et d'une grammaire de base*. Paris: Didier.
- Guerra Salas, L., Paredes García, F. y Gómez Sánchez, M.^a E. (2004). Disponibilidad léxica en la Comunidad de Madrid. XXXIV *Simpósio de la Sociedad Española de Lingüística*, Madrid, 13-16 de diciembre. Madrid: CSIC.
- Guillén de Rezzano, C. (1958). *Los centros de interés en la escuela* (5.^a ed.). Buenos Aires: Editorial Losada, S. A.
- Guillén Díaz, C. (2000). La enseñanza/aprendizaje de lenguas en la educación del siglo XXI: Un 'valor añadido' para el acceso al conocimiento. *Puerta a la lectura*, 9-10, 64-69. Disponible en <http://dialnet.unirioja.es/servlet/articulo?codigo=206246>
- Guntermann, G. (1995). The Peace Corps Experience: Language Learning in Training and in the Field. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 123-148). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hamers, J. F. y Blanc, M. (1982). Towards a social-psychological model of bilingual development. *Journal of Language and Social Psychology*, 1(1), 29-49.

- Hamers, J. F. y Blanc, M. (1989). *Bilinguality and Bilingualism*. Cambridge: Cambridge University Press.
- Henmon, V. A. C. (1924). *A French Word Book Based on a Count of 400.000 running words*. Madison, WI: Bureau of Educational Research, University of Wisconsin.
- Hernández Campoy, J. (1999). *Geolingüística: modelos de interpretación geográfica para lingüistas*. Murcia: Editum.
- Hernández Campoy, J. M. y Almeida, M. (2005). *Metodología de la investigación sociolingüística*. Granada: Comares.
- Hernández Muñoz, N. (2004). *El léxico disponible de los estudiantes conquenses*. Salamanca: Universidad de Salamanca.
- Hernández Muñoz, N. (2006). *Hacia una teoría cognitiva integrada de la disponibilidad léxica. El léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad de Salamanca.
- Hernández Solís, M.^a M. B. (2000). *Disponibilidad léxica de estudiantes de primaria de la ciudad de Zacatecas*. (Tesis de maestría inédita). México: Universidad Nacional Autónoma de México.
- Hernández Solís, M.^a M. B. (2006). *Un modelo para la planificación de la enseñanza del vocabulario con fundamento en el léxico disponible de universitarios zacatecanos*. (Tesis doctoral inédita). México: Universidad Nacional Autónoma de México.
- Hoffmann, E. T. A. (1887). *Le tonnelier de Nuremberg*. Traduction française par L. Jeanneret et E. Malvoisin. Paris: Hachette.
- Horst, M. (2005) Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61, 355-382.
- Howard, M. (2004). L'acquisition de la liaison en français langue seconde. Une analyse quantitative d'apprenants avancés en milieu guidé et en milieu naturel. *Cognition, Représentation, Langage*, 3(1),1-20.
- Howard, M. (2006). The expression of number and person through verb morphology in French interlanguage. *International Review of Applied Linguistics*, 4(1), 1-22.
- Howard, M. (2009). Short- versus long-term effects of naturalistic exposure on the advanced instructed learner's L2 development: A case study. En E. Labeau y

- F. Myles (Eds.), *The advanced learner variety: The case of French* (pp. 93-123). Oxford: Peter Lang.
- Hsu, H. L. (2013). Effects of Bilingualism and Trilingualism in L2 Production: Evidence from Errors and Self-Repairs in Early Balanced Bilingual and Trilingual Adults. *Journal of psycholinguistic research*, 43(4), 357-379.
- Huebner, T. (1995). The Effects of Overseas Language Programs: Report on a Case Study of an Intensive Japanese Course. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 149-170). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hugo, N. M. (2003). *Disponibilidad léxica en español como lengua extranjera* (trabajo de investigación inédito). Saarbrücken: Universität des Saarlandes.
- Ife, A., Vives Boix, G. y Meara, P. (2000). The impact of the study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, 4(1), 55-84.
- Instituto Cervantes. (2006). *Plan Curricular del Instituto Cervantes (PCIC)*. Madrid: Biblioteca Nueva. Disponible en http://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/default.htm
- Jacobsen, M. e Imhoof, M. (1974). Predicting success in learning a second language. *The Modern Language Journal*, 58(7), 329-336.
- Jessner, U. (1999). Metalinguistic Awareness in Multilinguals: Cognitive Aspects of Third Language Learning. *Language Awareness*, 8(3-4), 201-209.
- Jiménez Berrio, F. (2012a). Aprendizaje integrado de contenidos y español para inmigrantes desde la perspectiva de la disponibilidad léxica: enseñanza del vocabulario. En R. Breeze, F. Jiménez Berrio, C. Llamas Saíz, C. Martínez Pasamar y C. Tabernero Sala (Eds.), *Teaching approaches to CLIL / Propuestas docentes en AICLE* (pp. 149-163). Pamplona: Servicio de Publicaciones de la Universidad de Navarra. Disponible en <http://dadun.unav.edu/bitstream/10171/27574/1/Jim%C3%A9nez.pdf>

- Jiménez Berrio, F. (2012b). Disponibilidad léxica y competencia léxica en inmigrantes aprendices de español en Secundaria: los programas de inmersión lingüística. *Interlingüística*, 22, 35-45.
- Jiménez Berrio, Felipe (2013). *Léxico disponible de inmigrantes escolares no hispanohablantes*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. Disponible en: <http://dspace.unav.es/dspace/handle/10171/27719>
- Jiménez Catalán, R. (Ed.). (2014). *Lexical Availability in English and Spanish as a Second Language*. Dordrech: Springer Netherlands.
- Jing, L. (2006). *El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera*. (Memoria de Master inédita). Universidad de Alcalá.
- Johnson, J. S. y Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.
- Juilland, A., Brodin, D. y Davidovitch, C. (1970). *Frequency dictionary of French Words, the Romance Languages and their Structures, First Series, F1*. The Hague: Mouton.
- Justo Hernández, H. G. (1986). *La disponibilidad léxica en colores*. (Tesina inédita). Universidad Nacional Autónoma de México, México D. F.
- Kassaiian, Z. y Esmæ'li, S. (2011). The effect of bilinguality on L3 breadth of vocabulary knowledge and word reading skill. *Theory and Practice in Language Studies*, 1(8), 966-974.
- Katz, J. (1972). *Semantic Theory*. New York: Harper & Row.
- Katz, J. (1979). *Teoría Semántica* (J. García Puente, trad.). Madrid: Editorial Aguilar. (Obra original publicada en 1972).
- Katz, J. y Fodor J. A. (1963). The structure of a semantic theory. *Language*, 39, 170-210.
- Keshavarz, M. H. y Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language (L3). *Bilingual Education and Bilingualism*, 7(4), 295-302.
- Kinginger, C., y Farrell, K. (2004). Assessing Development of Meta-Pragmatic Awareness in Study Abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 19-42.

- Klein, W. (1986). *Second Language Acquisition*. Cambridge: Cambridge University Press.
- Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. London: Logman.
- Krashen, S. D. (1987). *Principles and Practice in Second Language Acquisition*. Hertfordshire: Prentice Hall International Ltd.
- Kripke, S. (1972). Naming and necessity. En D. Davidson y G. Harman (Eds.), *Semantics of Natural Language* (pp. 253-355). Dordrecht: Reidel.
- Kuhl, P. K. (2004). Early Language Acquisition: Cracking the Speech Code. *Nature Reviews Neuroscience*, 5(11), 831-843.
- Kuhl, P. K., Tsao, F. M., y Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100(15), 9096-9101.
- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, D. C.: Center for Applied Linguistics.
- Lafford, B. A. (1995). A Comparison of Communicative Strategies Used by Students Studying Spanish Abroad and 'At Home'. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 97-122). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Lagüéns Gracia, J. A. (2008). La variable sexo en el léxico disponible de los jóvenes aragoneses. En M.^a L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 103-162). Zaragoza: Institución Fernando el Católico.
- Lagüéns Gracia, V. (2008). La variable sexo en el léxico disponible de los jóvenes aragoneses. En M.^a L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 102-161). Zaragoza: Institución Fernando el Católico.
- Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind*. Chicago-London: The University of Chicago Press.
- Lapkin, S., Hart, D. y Swain, M. (1995). A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. En B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 67-94). Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Laroche, S., Richard, S. y Soulières, I. (2000). What some effects might not be: the time to verify membership in «well-defined» categories. *The Quarterly Journal of Experimental Psychology*, 53A(4), 929-961.
- Laroche, S., Richard, S. y Soulières, I. (2000). What some effects might not be: the time to verify membership in «well-defined» categories. *The Quarterly Journal of Experimental Psychology*, 53A(4), 929-961.
- Larsen-Freeman, D. y Long, M. H. (1991). *An Introduction to Second Language Acquisition Research*. London-New York: Longman.
- Le Play, F. (1855). *Les ouvriers européens, études sur les travaux, la vie domestique et la condition morale des populations ouvrières de l'Europe, précédées d'un exposé de la méthode d'observation*. Paris: Imprimerie impériale.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.
- Lerea, L. y Kohut, S. (1961). A comparative study of monolinguals and bilinguals in a verbal task performance. *Journal of clinical psychology*, 17(1), 49-52.
- Lévy Mangin, J. P. y Oubiña Barbolla, J. (2006). Análisis PATH y modelización con variables observables. En J. Varela Mallou (Dir.) y J. P. Lévy Mangin (Coord.), *Modelización con estructuras de covarianzas en ciencias sociales: temas esenciales, avanzados y aportaciones especiales* (pp. 175-191). La Coruña: Netbiblo.
- Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa. BOE núm. 187, de 6 de agosto de 1970, pp. 12525-12546. Disponible en <http://www.boe.es/boe/dias/1970/08/06/pdfs/A12525-12546.pdf>
- Lightbown, P. y Spada, N. M. (2006). *How Languages are learned* (3rd ed.). Oxford: Oxford University Press.
- Llanes Baró, À. (2010). *Children and adults learning English in a study abroad context*. (Tesis doctoral inédita). Universidad de Barcelona.
- Llanes Baró, À. (2011). The many faces of study abroad: an update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism*, 8(3), 189-215.

- Llanes Baró, À. y Muñoz Lahoz, C. (2009). A short stay abroad: does it make a difference? *System*, 37(3), 353-365.
- Llanes Baró, À. y Serrano Serrano, R. (2011). Length of stay and study abroad: Language gains in two versus three months abroad. *Revista española de lingüística aplicada*, 24, 95-110.
- Llopis Rodrigo, F. (2005). Transferències lèxiques en una comunitat de llengües en contacte. *Lenguaje y textos*, 23, 59-68.
- Llopis Rodrigo, F. (2008). *El lèxic del valencià central: un estudi de disponibilitat lèxica*. Valencia: Universidad de Valencia.
- Llopis Rodrigo, F. y Gómez Devís, M.^a B. (2010a). *El lèxic disponible de València*. Valencia: Editorial Denes.
- Llorente Pinto, M.^a R. (2005). Léxico disponible y léxico dialectal de la provincia de Ávila. En L. Santos Río, J. Borrego Nieto, J. F. García Santos, J. J. Gómez Asencio y E. Prieto de los Mozos (Eds.) y L. Santos Río (Coord.), *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter* (pp. 681-694). Salamanca: Ediciones Universidad de Salamanca.
- Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251-285.
- López Chávez, J. (1992). Alcances panhispánicos del léxico disponible. *Lingüística*, 4, 26-124.
- López Chávez, J. (1993). *El léxico disponible de escolares mexicanos*. México: Editorial Alhambra Mexicana.
- López Chávez, J. (1994). Comportamiento sintáctico de algunos verbos ordenados según su grado de disponibilidad léxica. *Revista de Estudios de Adquisición de la Lengua Española*, 1, 67-84.
- López Chávez, J. (1995). Léxico fundamental panhispánico: realidad o utopía. El español de América. En A. H. Matus Olivier *et al.* (Eds.), *Actas del IV Congreso Internacional de 'El Español de América'* (vol. 2, pp. 1006-1014). Santiago de Chile: Pontificia Universidad Católica de Chile.
- López Chávez, J. y Strassburger Frías, C. (1987). Otro cálculo del índice de disponibilidad léxica. *Presente y perspectivas de la investigación computacional en*

- México. *Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada*. México: Universidad Nacional Autónoma de México.
- López Chávez, J. y Strassburger Frías, C. (1991). Un modelo para el cálculo del índice de disponibilidad léxica individual. En H. López Morales (Ed.), *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre «Aportes de la lingüística a la enseñanza del español como lengua materna»* (pp. 91-112). Río Piedras: Universidad de Puerto Rico.
- López González, A. M. (2014). La estructura interna del léxico disponible en español como lengua extranjera (ELE) de los preuniversitarios polacos. *Studia Romanica Posnaniensia*, 41(1), 45-61. Disponible en <https://repozytorium.amu.edu.pl/jspui/bitstream/10593/11160/1/04%20srp41-1%20Antonio%20Mar%C3%ADa%20L%C3%B3pez%20Gonz%C3%A1lez.pdf>
- López Meirama, B. (2008). *Léxico disponible en el español de Galicia*. Santiago de Compostela: Universidad de Santiago de Compostela.
- López Meirama, B. (Ed.). (2011). *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidade de Santiago de Compostela.
- López Morales, H. (1973). *Disponibilidad léxica en escolares de San Juan*. MS.
- López Morales, H. (1978). Frecuencia léxica, disponibilidad y programación curricular. En H. López Morales (Ed.), *Aportes de la lingüística a la enseñanza del español como lengua materna*, número especial del *Boletín de la Academia Puertorriqueña de la Lengua Española*, 6(1), 73-86.
- López Morales, H. (1979). Disponibilidad léxica y estratificación socioeconómica. En H. López Morales (Ed.), *Dialectología y sociolingüística. Temas puertorriqueños* (pp. 173-181). Madrid, Miami, New York, San Juan: Hispanova de Ediciones.
- López Morales, H. (1983). Lingüística estadística. En H. López Morales (Coord.), *Introducción a la lingüística actual* (pp. 209-225). Madrid: Playor.
- López Morales, H. (1986). *Enseñanza de la lengua materna. Lingüística para maestros de español*. Madrid: Playor.

- López Morales, H. (1991). *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre Aportes de la Lingüística del Español como Lengua Materna*. Río Piedras: Universidad de Puerto Rico.
- López Morales, H. (1994). *Métodos de investigación lingüística*. Salamanca: Ediciones Colegio de España.
- López Morales, H. (1995-1996). Los estudios de disponibilidad léxica: pasado y presente. *Boletín de Filología de la Universidad de Chile, Homenaje a Rodolfo Oroz*, 35, 245–259.
- López Morales, H. (1999a). *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.
- López Morales, H. (1999b). Anglicismos en el léxico disponible de Puerto Rico. En L. A. Ortiz López (Ed.), *El Caribe Hispánico: perspectivas lingüística actuales. Homenaje a Manuel Álvarez Nazario* (pp. 147–170). Frankfurt am Main, Madrid: Iberoamericana Vervuert.
- López Morales, H. (2008). Un alto en el camino: el léxico disponible de Aragón. En M.^a L. Arnal Purroy, M.^a Luisa (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 7–15). Zaragoza: Institución Fernando el Católico. Disponible en http://ifc.dpz.es/recursos/publicaciones/27/27/_ebook.pdf
- López Morales, H. y García Marcos, F. J. (1995). Disponibilidad léxica en Andalucía. Proyecto de investigación. *Revista de Estudios de Adquisición de la Lengua Española*, 3, 65-76.
- López Rivero, E. (2008). *Estudio de disponibilidad léxica en 43 estudiantes de ELE*. (Memoria de máster). Universidad Antonio de Nebrija, Madrid. Disponible en http://www.mecd.gob.es/dctm/redele/Material-RedEle/Biblioteca/2008_BV_09/2008_BV_09_2_semestre/2008_BV_09_16Lopez_Rivero.pdf?documentId=0901e72b80e2ad5f
- Lord, G. (2010). The combined effects of immersion and instruction in second language pronunciation. *Foreign Language Annals*, 43(3), 488-503.
- Luque Toro, L. (Ed.). (2006). *Léxico español actual*. Venecia-Treviso: Cafoscarina.
- Lyster, R. (2007). *Learning and Teaching Languages through Content: A counterbalanced*. Amsterdam/Philadelphia: John Benjamins.

- Mackey, W. C. (1971a). *Le vocabulaire disponible du français* (2 vols.). Paris, Bruxelles, Montreal: Didier.
- Mackey, W. C. (1971b). *Le sondage dans les enquêtes de disponibilité*. Québec: Centre International de Recherche sur le Bilinguisme.
- Macnamara, J. (1969). How can one measure the extent of a person's bilingual proficiency? En L. G. Kelly (Ed.), *Description and measurement of bilingualism: An international seminar* (pp. 80-119). Toronto: University of Toronto Press.
- Maghsoudi, M. (2008). Learning English as a third language: a comparative study between Iranian and Indian bilinguals. *South Asian Language Review*, 28(1), 28-41.
- Mägisté, E. (1984). Learning a third language. *Journal of Multilingual and Multicultural Development*, 5(5), 415-421.
- Magnúsdóttir, S. (2012). *Disponibilidad léxica en alumnos de español como lengua extranjera. Estudio sobre el léxico disponible en alumnos de ELE en la secundaria en Islandia*. (Tesis de máster inédita). Sigillum Universitatis Islandiae. Disponible en http://skemman.is/stream/get/1946/12744/31050/3/Sigr%C3%BAn_Magn%C3%BAsd%C3%B3ttir.pdf
- Mangado, M. y Areta Lara, M. (2008). Procesamiento informático de datos para la elaboración de diccionarios de disponibilidad léxica. En I. Olza Moreno, M. Casado Velarde y R. González Ruiz, (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)* (pp. 479-494). Pamplona: Universidad de Navarra.
- Marriott, H. (1995). The Acquisition of Politeness Patterns by Exchange Students in Japan. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*, pp. 171-194. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Marslen-Wilson, W. D. (1989). Access and integration: Projecting sound onto meaning. En W. D. Marslen-Wilson (Ed.), *Lexical Representation and Process*. Cambridge, MA: The MIT Press.
- Martínez Olmos, E. (2008). *Disponibilidad léxica en las comunidades de habla alicantinas*. Alicante: Universidad de Alicante. Disponible en http://rua.ua.es/dspace/bitstream/10045/14307/1/tesis_martinez.pdf

- Mateo García, M.^a V. (1995). Disponibilidad léxica: posibles aplicaciones. En J. de D. Luque Durán y A. Pamies Bertrán (Eds.), *Actas de las I Jornadas sobre el estudio y la enseñanza del léxico* (pp. 141-147). Granada: Universidad de Granada.
- Mateo García, M.^a V. (1998). *Disponibilidad léxica en el COU almeriense: estudios de estratificación social*. Almería: Universidad de Almería.
- Maxwell, J. A. (1996). *Qualitative Research Design: An Interactive Approach*. London: Sage Publications.
- Maxwell, J. A. (2009). Designing a qualitative study (2nd ed.). En L. Bickman y D. J. Rog (Eds.), *The SAGE Handbook of Applied Social Research Methods* (pp. 214-246). Los Ángeles: Sage.
- McClellan, J. y Rumelhart, D. (1986). *Parallel Distributed Processing: Explorations in the microstructure of Cognition*. Cambridge, MA: The MIT Press.
- Mena Osorio, M. (1986). *Disponibilidad léxica infantil en tres niveles de enseñanza básica* (Tesis de maestría inédita). Universidad de Concepción.
- Michéa, R. (1949). Introduction pratique a une statistique du langage. *Les Langues Modernes*, 43(3A), 173-186.
- Michéa, R. (1950). Vocabulaire et culture. *Les Langues Modernes*, 44, fascicule B, 188-189.
- Michéa, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les Langues Modernes*, 47, 338-344.
- Michéa, R. (1959). *Vocabulaire allemand progressif*. Paris: Didier.
- Miller, L. y Ginsberg, R. B. (1995). Folklinguistic Theories of Language Learning. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 293-316). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Milroy, L. (1987). *Observing and analysing natural language: a critical account of sociolinguistic method*. *Language in society*, 12. Oxford: Basil Blackwell.
- Milton, J. y Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *Review of Applied Linguistics*. 107-108, 17-34.
- Moliné Juste, A. B. (2008). Creación léxica en el vocabulario disponible de los jóvenes aragoneses: aspectos morfológicos. En M.^a L. Arnal Purroy (Ed.),

- Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 163-193). Zaragoza: Institución Fernando el Católico.
- Montenegro, L. y Haché de Yunén, A. (1997). Una propuesta para el desarrollo de estrategias de comprensión lectora: Tiempo de Lectura. En M. C. Martínez (Comp.), *Los procesos de la lectura y la escritura*. Santiago de Cali: Editorial Universidad del Valle.
- Montenegro, L. y Haché de Yunén, A. (2001). *Tiempo de Lectura*. Israel: Edusoft Ltd. PUCMM y CET.
- Montenegro, R. (2010). *Estudio de disponibilidad léxica en estudiantes de tercero y sexto de primaria de escuelas públicas del departamento de Guatemala*. (Informe breve de investigación, 1). Guatemala: United States Agency International Development. Disponible en <http://www.reaula.org/administrador/files/Infor.%20breves%20de%20inves%20tiga%20No.1%20fc.pdf>
- Moon, J. (2000). *Children Learning English*. Oxford: MacMillan-Heinemann.
- Moreno Fernández, F. (1990). *Metodología sociolingüística*. Madrid: Gredos.
- Moreno Fernández, F. (2009). *Principios de sociolingüística y sociología del lenguaje* (4.^a ed.). Barcelona: Ariel.
- Moreno Fernández, F. (2012). Disponibilidad léxica: cuestiones metodológicas: A propósito de disponibilidad léxica de los estudiantes hispanos de Redwood City, CA. *Revista Nebrija de Lingüística Aplicada*, 11 [en línea]. Disponible en http://nebrija.com/revista-linguistica/files/articulosPDF/articulo_5285da9680d3e.pdf
- Moreno Fernández, F., Moreno Fernández, J. E. y García de las Heras, A. J. (1995). Cálculo de disponibilidad léxica. El programa Lexidisp. *Lingüística*, 7, 243–249.
- Morgan, G. A. y Griego, O. V. (1998). *Easy Use and Interpretation of SPSS for Windows: Answering Research Questions With Statistic*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mougeon, R., Nadasdi, T., y Rehner, K. (2010). *The sociolinguistic competence of immersion students*. Bristol: Multilingual Matters.

- Müller, C. (1963). Le Mot, unité de texte et unité de lexique en statistique lexicologique. *Travaux de linguistique et de littérature*, 2(1), 155-175.
- Müller, C. (1969). La statistique lexicale. *Langue française*, 2, 30-43.
- Muñoz Lahoz, C. (2002). *Aprender idiomas*. Barcelona: Ediciones Paidós.
- Murillo Rojas, M. (1991). *Disponibilidad léxica en los preescolares del kínder María Jiménez, según sexo, ocupación del padre y lugar de procedencia del niño: fundamentos básicos para una planificación lingüística*. (Tesis de licenciatura). Universidad de Costa Rica.
- Murillo Rojas, M. (1993). Disponibilidad léxica en los preescolares: estudio de cinco campos semánticos. *Káñina, Revista de Artes y Letras de la Universidad de Costa Rica*, 17(2), 117-127.
- Murillo Rojas, M. (1994). Comidas y bebidas: estudio de la disponibilidad léxica en preescolares. *Káñina, Revista de Artes y Letras de la Universidad de Costa Rica*, 18(2), 117-133.
- Nayak, H., Hansen, N., Krueger, N. y McLaughlin, B. (1990). Language-Learning Strategies in Monolingual and Multilingual Adults. *Language Learning*, 40(2), 221-244.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, 14, 11-28.
- Njock, P. E. (1978). *L'univers familier de l'enfant africain*. Québec: Centre International de Recherche sur le Bilinguisme. Disponible en <http://0-files.eric.ed.gov.opac.msmc.edu/fulltext/ED192580.pdf>
- Ó Huallachain, C. L. (1966). *Buntus Gaeilge*. Vaile Atha Cliath, Dublin: Oifig Dhiolta Foilseachan Rialtais (Government Publications Sales Offices), An Stuara Ard-Oifig an Phoist (The Arcade, General Post-Office).
- Ó Huallachain, C. L. (1974). *The Study of Vocabulary Availability in Ireland*. Fourth Annual TESOL Conference, del 18 a 21 de marzo en San Francisco. Disponible en <http://files.eric.ed.gov/fulltext/ED041256.pdf>
- Ó Huallachain, C. L. (1978). Languages of Instruction in Ireland 1904-1977. *International Review of Education*, 24(4), 501-510.
- O'Grady, W. (2005). *How Children Learn Language*. Cambridge: Cambridge University Press.

- Ogden, C. K. (1930). *Basic English, a general introduction with rules and grammar*. London: K. Paul, Trench, Trubner & Co.
- Ogden, C. K. (1931). *Brighter Basic; examples of Basic English for young persons of taste and feeling*. London: K. Paul, Trench, Trubner & Co.
- Ogden, C. K. (1933). *Basic by examples*, London: K. Paul, Trench, Trubner & Co.
- Ogden, C. K. (1935). *Basic step by step*, London: K. Paul, Trench, Trubner & Co.
- Ogden, C. K. (1968). *Basic English, international second language*. New York: Harcourt, Brace & World.
- Ogden, C. K. y Richards, I. A. (1928). *The Basic English*. London.
- Orden de 8 de julio de 1971, sobre actividades docentes de los Institutos de Ciencias de la Educación en relación con la formación pedagógica de los universitarios. BOE núm. 192, de 12 de agosto de 1971, p. 13170. Disponible en <http://www.boe.es/boe/dias/1971/08/12/pdfs/A13170-13170.pdf>
- Orden ECI/3858/2007, de 27 de diciembre, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de las profesiones de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas. BOE núm. 312, de 29 de diciembre de 2007, pp. 53751-53753. Disponible en <http://www.boe.es/boe/dias/2007/12/29/pdfs/A53751-53753.pdf>
- Osorio, J. y Véliz, M. (2001). Diseño de un software para el desarrollo de la capacidad de lectura crítica. *RLA. Revista de lingüística teórica y aplicada*, 39, 203-220.
- Paredes García, F. (1999). La ortografía en las encuestas de disponibilidad léxica. *Revista Estudio Adquisición de la Lengua Española (REALE)*, 11, 75-98.
- Paredes García, F. (2014). A vueltas con la selección de ‘centros de interés’ en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE [en línea]. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 16. Disponible en <http://www.nebrija.com/revista-linguistica/a-vueltas-con-la-seleccion-de-centros-de-interes-en-los-estudios-de-disponibilidad-lexica-para-una-propuesta-renovadora-a-proposito-de-la-disponibilidad-lexica-en-ele>

- Paredes García, F. (Coord.), Álvaro García, S., Pérez, A., Rebollo, M. y Ruiz, S. (2002). *Léxico disponible de los estudiantes de educación secundaria de Alcalá de Henares y su comarca [en prensa]*. Alcalá de Henares: Servicio de Publicaciones de Alcalá de Henares.
- Pejović, A., Sekulić, M. y Karanović, V. (2012). *Comida y bebida en la lengua española, cultura y literaturas hispánicas*. Kragujevac: Facultad de Filología y Artes, Universidad de Kragujevac.
- Pellegrino, V. (1998). Students' perspectives on language learning in a study abroad context. En B. Freed (Ed.), *Second Language Acquisition in a Study abroad Context*, 91–120. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Pendanx, M. (1998). *Les activités d'apprentissage en classe de langue*. Paris: Hachette F.L.E.
- Pérez Vidal, C., y Juan-Garau, M. (2009). The effect of Study Abroad (SA) on written performance. *Eurosla yearbook*, 9, 269-295.
- Pfeffer, J. A. (1964). *Grunddeutsch Basic (spoken) German Word List Grundstufe*. New Jersey: Prentice Hall Inc. Englewood Cliffs.
- Pigada, M. y Schmitt, N. (2006) Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Polanyi, L. (1995). Language Learning and Living Abroad: Stories from the Field. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*, pp. 271-292. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Popper, K. (1959). *The Logic of Scientific Discovery*. New York: Basic Books.
- Prado Aragonés, J., Galloso Camacho, M.^a V. y Conceição, M. C. (2009). *La disponibilidad léxica en situación de contacto de lenguas en las zonas limítrofes de Andalucía y Extremadura (España) y Algarve y Alentejo (Portugal)*. Huelva: Universidad de Huelva.
- Prado Aragonés, M.^a V. y Galloso Camacho, J. (2005). *Léxico disponible de Huelva. Nivel preuniversitario*. Huelva. Universidad de Huelva.
- Prado Aragonés, M.^a V. y Galloso Camacho, J. (2008). *Léxico disponible de Huelva. Nivel 6.º de Educación Primaria*. Huelva. Universidad de Huelva.
- Priyono, P. (2009). Towards a Reduction of Grammar Teaching a Lexical Analysis. *TEFLIN Journal: A publication on the teaching and learning of English*, 10(1), 1-15.

- Puga, J. (2001). Una presentación del software multimedial Actos de Habla. *RLA. Revista de lingüística teórica y aplicada*, 39, 139-154.
- Puren, C. (1998). Didactique scolaire des langues vivantes étrangères en France et didactique française du français langue étrangère. *ÉLA. Études de Linguistique Appliquée*, 111, 359-383.
- Putnam, H. (1975). The meaning of 'meaning'. En K. Gunderson (Ed.), *Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science*, 7 (pp. 131-193). Minneapolis: University of Minnesota Press.
- Quilllian, M. R. (1968). Semantic memory. En M. Minsky (Ed.), *Semantic Information Processing* (pp. 227-70). Cambridge, MA: The MIT Press.
- Real Academia Española y Asociación de Academias de la Lengua Española (2005). *Diccionario panhispánico de dudas*. Madrid: Santillana. Disponible en <http://www.rae.es/recursos/diccionarios/dpd>
- Real Academia Española y Asociación de Academias de la Lengua Española (2010). *Ortografía de la lengua española*. Madrid: Espasa.
- Real Academia Española. (2001). *Diccionario de la lengua española* (22.^a ed.). Madrid: Editorial Espasa Calpe. Disponible en <http://www.rae.es/rae.html>
- Real Academia Española. (2006). *Diccionario esencial de la lengua española*. Madrid: Espasa Calpe. Disponible en <http://www.rae.es/recursos/diccionarios/desen>
- Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales. BOE núm. 260 de 30 de octubre de 2007, pp. 44037-44048. Disponible en <http://www.boe.es/boe/dias/2007/10/30/pdfs/A44037-44048.pdf>
- Real Decreto 1591/2010, de 26 de noviembre, por el que se aprueba la Clasificación Nacional de Ocupaciones 2011. BOE núm. 306, de 17 de diciembre de 2010, Sección I, pp. 104040-104060. Disponible en <http://www.boe.es/boe/dias/2010/12/17/pdfs/BOE-A-2010-19389.pdf>
- Real Decreto 850/1993, de 4 de junio, por el que se regula el ingreso y la adquisición de especialidades en los Cuerpos de Funcionarios Docentes a que se refiere la Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo. BOE núm. 155 de 30 de junio de 1993, pp. 19924-19941.

- Disponible en <http://www.boe.es/boe/dias/1993/06/30/pdfs/A19924-19941.pdf>
- Real Decreto 861/2010, de 2 de julio, por el que se modifica el Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales. BOE núm. 161, de 3 de julio de 2010, pp. 58454-58468. Disponible en <http://www.boe.es/boe/dias/2010/07/03/pdfs/BOE-A-2010-10542.pdf>
- Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. BOE núm. 35 de 29 de enero de 2011, pp. 13909-13926. Disponible en <http://www.boe.es/boe/dias/2011/02/10/pdfs/BOE-A-2011-2541.pdf>
- Rebagliato M., Ruiz I. y Arranz M. (1996). *Metodología de investigación en Epidemiología*. Madrid: Díaz de Santos.
- Regan, V. (1995). The Acquisition of Sociolinguistic Native Speech Norms: Effects of a Year Abroad on Second Language Learners of French. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 245-268). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Regan, V., Howard, M. y Lemée, I. (2009). *The Acquisition of Sociolinguistic Competence in a Study Abroad Context*. Bristol: Multilingual Matters.
- Richards, I. A. (1935). *Basic in teaching, East and West*. London: K. Paul, Trench, Trubner & Co.
- Richards, I. A. (1943). *Basic English and its uses*. London: K. Paul, Trench, Trubner & Co.
- Richards, I. A. (1945). *The pocket book of basic English, a self-teaching way into English*. New York: Pocket Books.
- Richards, J. C. (1969). A Psycholinguistic Measure of Vocabulary Selection. *International Review of Applied Linguistics in Language Teaching*, 8(2), 87-102. Disponible en <http://files.eric.ed.gov/fulltext/ED035860.pdf>
- Rivenc, P. (1956). Réponse au compte-rendu fait de «*Français élémentaire? Non*» par Roger G. *Europe*, 121-122.
- Rodríguez Bou, I. (Dir.). (1952). *Recuento de vocabulario español*. San Juan de Puerto

Rico: Organización de Estados Americanos.

Rodríguez Cadena, Y. (1997). *Disponibilidad léxica en los estudiantes de educación media en Barranquilla*. (Trabajo de conclusión de curso de pregrado inédito). Colombia: Universidad Del Atlántico.

Rodríguez Gómez, G., Gil Flores, J. y García Jiménez, E. (1996). *Metodología de la investigación cualitativa*. Málaga: Ediciones Algibe.

Rodriguez, L. (1993). Disponibilité lexicale chez les jeunes Franco-Manitobains en 1990-1991: méthode et résultats partiels. En H. J. Niedereche y L. Wolf, *Français du Canada-Français de France* (pp. 163-178). Tübingen: Max Niemeyer Verlag.

Rodriguez, L. (2006). *La langue française au Manitoba (Canada): histoire et évolution lexicométrique* (Vol. 21). Tübingen: Max Niemeyer Verlag.

Rodriguez, L. (2007). Ruralité et acquisition lexicale au Manitoba: le vocabulaire disponible dans les écoles Saint-Eustache (milieu rural) et Provencher (milieu urbain). *Cahiers franco-canadiens de l'Ouest*, 19(2), 141-158.

Rodriguez, L. (2009). De la langue au discours: évaluation lexicométrique de l'impact anglais sur le français en usage au Manitoba (Canada). En B. Pöll y E. Schafroth (Ed.), *Normes et hybridation linguistiques en francophonie. Actes de la section 6 du Congrès de l'Association des francoromanistes allemands, Augsburg, 24-26 septembre 2008* (pp. 157-182). Paris: L'Harmattan. Disponible en http://www.romanistik.hhu.de/fileadmin/redaktion/Fakultaeten/Philosophische_Fakultaet/Romanistik/Romanistik_4_Sprachwissenschaft/Dateien/Schafroth_Detailseite/Publikationen/normes_et_hybridation.pdf

Rodriguez, L. (2010). Le rôle du marquage lexicométrique des anglicismes dans un corpus franco-canadien. En S. Bolasco, I. Chiara I. y L. Guiliano L. (Ed.), *Jadt2010, Statistical Analysis of Textual Data. Proceedings of the 10th International Conference Journées d'Analyse statistique des Données Textuelles* (pp. 1295-1303). Rome: Sapienza, University of Rome. Disponible en http://lexicometrica.univ-paris3.fr/jadt/jadt2010/allegati/JADT-2010-1295-1304_181-Rodriguez.pdf

- Rojas Mayer, E. M. (Coord.) (2012). *Léxico e interculturalidad: nuevas perspectivas*. San Miguel de Tucumán: Instituto de Investigaciones Lingüísticas y Literarias Hispanoamericanas.
- Román, B. (1985). *Disponibilidad léxica en españoles de Dorado, Puerto Rico*. (Memoria de licenciatura inédita). Universidad de Puerto Rico.
- Romero Rubilar, C. (2000). Variabilidad y prototipicidad en vocabulario disponible. *RLA. Revista de lingüística teórica y aplicada*, 38, 139-148.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. En T. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 111-144). New York: Academic Press.
- Rosch, E. (1978). Principles of categorization. En E. Rosch y B. Lloyd, *Cognition and categorization* (pp. 24-48). Hillsdale, NJ: Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. y Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology*, 8, 382-439. Disponible en <http://www.cns.nyu.edu/~msl/courses/2223/Readings/Rosch-CogPsych1976.pdf>
- Rueda Bernao, M.^a J. (1994). Análisis de cuatro manuales para la enseñanza del español lengua extranjera. *Revista Estudio Adquisición de la Lengua Española (REALE)*, 2, 79-114.
- Ruiz Basto, A. (1987). *Disponibilidad léxica de los alumnos de primer ingreso en el Colegio de Ciencias y Humanidades Plantel Naucalpan*. (Memoria de licenciatura inédita). México D. F.: Universidad Autónoma de México.
- Ruiz Olabuénaga, J. I. (1999). *Metodología de la Investigación Cualitativa* (5.^a ed.). Bilbao: Universidad de Deusto Editores.
- Ruiz Olabuénaga, J. I. (2003). *Técnicas de triangulación y control de calidad en la investigación socioeducativa*. Bilbao: Ediciones Mensajero.
- Rumelhart, D., Lindsay, P. y Norman, D. (1972). A process model for long-term memory. En E. Tulving y W. Donaldson, *Organization and Memory* (pp. 197-246). New York: Academic Press.
- Saine Camargo, A. M.^a (2008). *El léxico disponible de estudiantes de la escuela media bonaerense. Aspectos metodológicos y sociolingüísticos*. (Tesis doctoral inédita).

Universidad Nacional de Educación a Distancia.

Salkind, N. J. (1998). *Métodos de investigación*. México: Prentice Hall.

Samper Hernández, M. (2002). *Disponibilidad léxica en alumnos de español como lengua extranjera*. Colección Monografías n.º 4. Málaga: ASELE.

Samper Hernández, M. (2004). Principios generales y bases metodológicas para el estudio de la evolución de la disponibilidad léxica en escolares. En C. Díaz Alayón y M. Morera Pérez (Coords.), *Homenaje a Francisco Navarro Artilles*, pp. 453-474. Madrid: Academia Canaria de la Lengua.

Samper Hernández, M. (2005). Dialectalismos en el léxico disponible de escolares grancanarios. En L. Santos Río (Coord.), *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter*, pp. 1065-1078. Salamanca: Universidad de Salamanca.

Samper Hernández, M. (2008). Datos comparativos entre dos léxicos disponibles: Aragón y Gran Canaria. En M.^a L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 245-282). Zaragoza: Institución Fernando el Católico.

Samper Hernández, M. (2009). *Evolución de la disponibilidad léxica en estudiantes grancanarios de enseñanza primaria y secundaria*. Las Palmas de Gran Canaria: Cabildo de Gran Canaria.

Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible: sugerencias. *Lingüística*, 10, 311-33.

Samper Padilla, J. A. (1999). Léxico disponible y variación dialectal: datos de Puerto Rico y Gran Canaria. En E. Forastieri Braschi, J. Cardona, H. López Morales y A. Morales de Walters (Coords.), *Estudios de Lingüística Hispánica (Homenaje a María Vaquero)* (pp. 551-573). San Juan de Puerto Rico: Editorial de la Universidad de Puerto Rico.

Samper Padilla, J. A. (2003). El «Proyecto Panhispánico» de disponibilidad léxica: logros y estado actual. En M. Casas Gómez. M. (Dir.) y C. Varo Varo (Coord.), *VII Jornadas de Lingüística* (pp. 193-225). Cádiz: Universidad de Cádiz.

Samper Padilla, J. A. y Hernández Cabrera, C. (1997). El estudio de la disponibilidad léxica en Gran Canaria: datos iniciales y variación sociolingüística. En M. Almeida y J. Dorta, *Contribuciones al estudio de la lingüística hispánica. Homenaje al*

- profesor Ramón Trujillo Carreño* (tomo II, pp. 229–239). La Laguna: Universidad de la Laguna.
- Samper Padilla, J. A. y Hernández Cabrera, C. E. (2006). Densidad de dialectalismos y condicionantes sociales en el léxico disponible de Canarias. En M. Sedano, A. Bolívar y M. Shiro, *Haciendo lingüística. Homenaje a Paola Bentivoglio* (pp. 537-553). Caracas: Universidad Central de Venezuela.
- Samper Padilla, J. A., Bellón Fernández, J. J. y Samper Hernández, M. (2003). El proyecto de estudio de la disponibilidad léxica en español En R. Ávila, Samper Padilla e H. Ueda (Eds.), *Pautas y pistas en el análisis del léxico hispano(Americano)* (pp. 27-140). Madrid/Frankfurt am Main: Iberoamericana Vervuert.
- Samper Padilla, J. A., Hernández Cabrera, C. E. y Bellón Fernández, J. J. (2003). Léxico disponible y variación dialectal: datos de Canarias y Córdoba. En C. Díaz Alayón, M. Morera y G. Ortega (Eds.), *Estudios sobre el español de Canarias. Actas del I Congreso Internacional sobre el español de Canarias* (vol. II, pp. 1077-1099). Las Palmas de Gran Canaria: Academia Canaria de la Lengua.
- Sánchez Corrales, M. y Murillo Rojas, V. M. (2005). Elementos diaintegrativos del léxico básico escolar costarricense. *Kánina, Revista de Artes y Letras de la Universidad de Costa Rica*, 29, 165-177.
- Sánchez Corrales, M. y Murillo Rojas, V. M. (2006). *Disponibilidad léxica de los niños preescolares costarricenses*. Costa Rica: Universidad de Costa Rica.
- Sánchez Corrales, M. y Murillo Rojas, V. M. (1993). Campos semánticos y disponibilidad léxica en preescolares. *Revista de Educación*, 17(2), 15-25.
- Sánchez-Saus Laserna, M. (2008). *Léxico y enseñanza del español. Disponibilidad léxica de los estudiantes de español como lengua extranjera de la Universidad de Cádiz* (trabajo de investigación). Universidad de Cádiz.
- Sánchez-Saus Laserna, M. (2011). *Bases semánticas para el estudio de los centros de interés del léxico disponible. Disponibilidad léxica de informantes extranjeros en las universidades andaluzas*. (Tesis doctoral). Universidad de Cádiz. Disponible en <http://rodin.uca.es/xmlui/handle/10498/15862>

- Sankoff, G. (1974). A quantitative paradigm for the study of communicative competence. En R. Bauman y J. Sherzer (Eds.) *Explorations in the Ethnography of Speaking* (pp. 18–49). Cambridge: Cambridge University Press.
- Sankoff, G. (Ed.). (1980). *The social life of language*. Philadelphia: University of Pennsylvania Press.
- Santiago Guervós, J. de (2008). *Léxico disponible de Segovia*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.
- Santos Díaz, I. C. (2013a). Bienvenido a tu casa: propuesta didáctica para el aprendizaje de español como segunda lengua. En L. Rodríguez García y A. R. Roldán Tapia (Coord.), *Relaciones Interculturales en la Diversidad* (pp. 167-176). Córdoba: Cátedra Intercultural de la Universidad de Córdoba.
- Santos Díaz, I. C. (2013b). Enseñanza de español como segunda lengua en el contexto escolar a través de AICLE. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 13 [en línea]. Disponible en http://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_532c3e084bcbf.pdf
- Santos Díaz, I. C. (2014a). La influencia del conocimiento de varias lenguas extranjeras en el léxico disponible. Comunicación presentada en el *XI Congreso Internacional de Lingüística General*, del 21 al 23 de mayo, Universidad de Navarra.
- Santos Díaz, I. C. (2014b). Los verbos de afección: dificultades en el aprendizaje de español como lengua extranjera. En C. J. Álvarez López, B. Garrido Martín y M. González Sanz, *Jóvenes aportaciones a la investigación lingüística* (pp. 657-670). Sevilla: Ediciones Alfar.
- Santos Díaz, I. C. (2015a). El impacto de las estancias en el extranjero en el léxico disponible en inglés y francés de estudiantes de postgrado. *E-AESLA*, 1 [en línea]. Disponible en <http://cvc.cervantes.es/lengua/eaesla/pdf/01/32.pdf>
- Santos Díaz, I. C. (2015b). Análisis comparativo del léxico en español y en lengua extranjera del futuro profesorado. *E-AESLA*, 1 [en línea]. Disponible en <http://cvc.cervantes.es/lengua/eaesla/pdf/01/67.pdf>

- Sanvicens Gila, A. M.^a (2007). *Aspectos teóricos para el estudio del léxico disponible en comunidades bilingües: el caso de Navarra*. (Trabajo de investigación de Tercer Ciclo). Universidad de Navarra.
- Saralegui, C. y Tabernero Sala, C. (2008). *Aportación al «Proyecto Panhispánico» de léxico disponible: Navarra*. En I. Olza Moreno, M. Casado Velarde y R. González Ruiz (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)* (pp. 746-761). Pamplona: Servicio de Publicaciones de la Universidad de Navarra.
- Sasaki, M. (2009). Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. En R. M. Machón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 49-76). Clevedon: Multilingual Matters.
- Saura Rami, J. (2008). La ortografía en las encuestas aragonesas de disponibilidad léxica. En M. L. (Ed.), *Estudios sobre disponibilidad léxica en jóvenes aragoneses* (pp. 195-206). Zaragoza: Institución Fernando el Católico.
- Sax, K. (2003). *Acquisition of stylistic variation by American learners of French*. (Unpublished PhD). Indiana University.
- Schank, R.C. (1975). *Conceptual Information Processing*. New York: Elsevier.
- Schreuder, R. y Flores D'Arcais, G. B. (1989). *Psycholinguistic issues in the lexical representation of meaning*. En W. Marslen-Wilson (Ed.) *Lexical Representation and Process* (pp. 409-436). Cambridge, MA: The MIT Press.
- Schlyter, B. (1951). *Centrala ordförrådet i franska*. Uppsala: Almqvist et Wiksell.
- Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20, 213-223.
- Searle, J. R. (1975). Indirect speech acts. En P. Cole y J. L. Morgan (Eds.), *Syntax and Semantics* (vol. 3. pp. 59-82). New York: Academic Press.
- Serrano Zapata, M. (2003). Disponibilidad léxica en la provincia de Lleida: aspectos metodológicos. *Interlingüística*, 12, 929-937.
- Serrano Zapata, M. (2004). Aspectos sociolingüísticos del léxico disponible castellano de los preuniversitarios leridanos. *Pragmalingüística*, 12, 147-165.

- Serrano Zapata, M. (2006). Consecuencias del contacto de lenguas en Lérida: interferencias detectadas en las encuestas de disponibilidad léxica. En J. L. Blas Arroyo, M. Casanova Ávalos y M. Velando Casanova (Eds.), *Discurso y sociedad: contribuciones al estudio de la lengua en contexto social* (pp. 811-829). Castellón de la Plana: Universidad Jaume I de Castellón.
- Serrano Zapata, M. (2014). *Disponibilidad léxica en la provincia de Lleida: estudio comparado de dos lenguas en contacto*. (Tesis doctoral inédita). Universitat de Lleida. Disponible en: <http://www.tdx.cbuc.es/handle/10803/285008>
- Siegal, M. (1995). Individual Differences and Study Abroad: Women Learning Japanese in Japan. En B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 197-224). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Šifrar Kalan, M. (2012). El léxico disponible de comida y bebida. En A. Pejović, M. Sekulić y V. Karanović (2012). *Comida y bebida en la lengua española, cultura y literaturas hispánicas* (pp. 191-207). Kragujevac: Facultad de Filología y Artes, Universidad de Kragujevac.
- Simões, A. (1996). Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania*, 79, 87-95.
- Singleton, D. (2001). Age and second language acquisition. *Annual Review of Applied Linguistics*, 21, 77-89.
- Spadaro, K. M. (1998). *Maturation constraints on lexical acquisition in a second language*. (Tesis doctoral inédita). University of Western Australia.
- Sperber, D. y Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
- Spolsky, B. (1969). Attitudinal aspects of second language learning. *Language Learning*, 19(3-4), 271-285.
- Stevens, J. (2001). Study abroad learners' acquisition of the Spanish voiceless stops. *MIFLC Review*, 10, 137-151.
- Stevens, S. S. (Ed.). (1946). *Handbook of Experimental Psychology*. New York: Wiley.
- Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. Cambridge: Cambridge University Press.

- Swain, M., Lapkin, S., Rowen, N. y Hart, D. (1990). The role of mother tongue literacy in third language learning. *Language, culture and curriculum*, 3(1), 65-81.
- Taber, C. R. y Nida, E. A. (1971). *La traducción: théorie et méthode*. Londres: Alliance Biblique Universelle.
- Tabernero Sala, C. (2008a). El léxico disponible como fuente de aproximación al estudio de los regionalismos. En I. Olza Moreno, M. Casado Velarde y R. González Ruiz, (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)* (pp. 811-824). Pamplona: Universidad de Navarra.
- Tabernero Sala, C. (2008b). Disponibilidad léxica y contacto de lenguas. *Oihenart. Cuadernos de lengua y literatura*, 23, 545-565.
- Taylor, S. J. y Bogdan, R. (1987). *Introducción a los métodos cualitativos* (2.ª ed.). Barcelona: Ediciones Paidós Ibérica, S. A.
- Terrell, T. (1986). Acquisition in the natural approach: The binding/access framework. *The Modern Language Journal*, 70, 213–227.
- Tharp, J. B. (1939). *The Basic French Vocabulary*. New York: Henry Holt.
- Thomas, J. (1992). Metalinguistic awareness in second- and third-language learning. *Advances in psychology*, 83, 531-545.
- Tomé Cornejo, C. (2011). Reflexiones en torno a los centros de interés». En M.ª V. Galloso Camacho (Coord.), *La investigación de la Lengua y la Literatura en la Onubense* (pp. 115-159). Huelva: Universidad de Huelva.
- Torres González, A. N. y García Falero, A. B. (2006). Planteamiento de la investigación de disponibilidad léxica en Tenerife. Comunicación presentada en el *III Congreso Internacional sobre lengua y sociedad*, Castellón de la Plana, 19-21 de septiembre .
- Trigo Ibáñez, E. y González Martínez, A. (2011). Estudio del comportamiento de la variable sexo en el léxico disponible de los preuniversitarios sevillanos. *Diálogo de la Lengua*, 3, 28-41
- Tulving, E. (1972). Episodic and semantic memory 1. En E. Tulving y W. Donaldson (Eds.), *Organization of Memory* (pp. 381-402). New York: Academic Press.

- Urrutia Martínez, M. A. (2001). *Edad y acceso al léxico*. (Tesis de magíster inédita). Universidad de la Concepción.
- Urzúa, P., Sáez, K. y Echeverría, M. S. (2006). Disponibilidad léxica matemática. Análisis cuantitativo y cualitativo. *RLA. Revista de Lingüística Teórica y Aplicada*, 44(2), 59-76. Disponible en http://www.scielo.cl/scielo.php?pid=S0718-48832006000200005&script=sci_arttext
- Valencia Espinoza, A. (1998). Léxico estudiantil. comentario sobre un repertorio. *Homenaje a Ambrosio Rabanales, Boletín de Filología de la Universidad de Chile*, 37, 1211-1221.
- Valencia Espinoza, A. (2000). Ciencia y tecnología. análisis de léxico estudiantil. *Nueva Revista del Pacífico*, 45, 89-98.
- Valencia Espinoza, A. y Echeverría, M. S. (1999) *Disponibilidad léxica en estudiantes de cuarto año de Educación Media*. Santiago de Chile: Universidad de Chile y Universidad de Concepción.
- Valencia Espinoza, A. y Echeverría, M. S. (1999). *Disponibilidad léxica en estudiantes chilenos*. Santiago de Chile: Universidad de Chile y Universidad de Concepción.
- Valencia Espinoza, A., y Echeverría, M. S. (1999). *Disponibilidad léxica en estudiantes chilenos*. Santiago de Chile: Universidad de Chile y Universidad de Concepción.
- Vander Beke, G. E. (1929). *French Word Book*, New York, The MacMillan Company. Disponible en <http://hdl.handle.net/2027/mdp.39015034357353>
- Vargas Sandoval, P. (1991). La disponibilidad léxica: un diagnóstico al hablante medio de la V región en seis centros de interés. *Nueva Revista del Pacífico*, 33-36, 115-123, .
- Véliz, M. (2001). Diseño de un software para el desarrollo de la capacidad de lectura crítica. *RLA. Revista de lingüística teórica y aplicada*, 39, 203-220.
- Véliz, M. (2002). Desarrollo de estrategias de lectura por medio del computador: evaluación de una experiencia. En G. Parodi (Ed.), *Lingüística e interdisciplinariedad: Desafíos del nuevo milenio. Ensayos en honor a Marianne Peronard* (pp. 161-178). Valparaíso: Ediciones Universitarias de Valparaíso UCV.
- Verlée, L. (1954). *Basis-Woordenboek voor de Franse*. Antwerpen, Amsterdam: De Sikkels, J. M. Meulenhoff.

- Viader, F., Lambert, J., de la Sayette, V., Eustache, F., Morin, P., Morin, I. et Lechevalier, B. (2002). Aphasie. Encyclopédie Médico-Chirurgicale. *Neurologie*, 17, 1-32. Disponible en <http://psychologie-m-fouchey.psyblogs.net/public/fichiers%20jointes/d%C3%A9mence/Aphasie.pdf>
- Victory, J. B. Jr. (1971). *A study of Lexical Availability among Monolingual-Bilingual Speakers of Spanish and English*. (Master Thesis). Houston (Texas): Rice University. Disponible en <http://hdl.handle.net/1911/27482>
- Vinay J. P. y Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais*. Paris: Marcel Didier.
- Vives Boix, G. (1995). *The development of a measure of lexical organisation: the Association Vocabulary Test*. (Tesis doctoral). University of Wales, Swansea.
- Waring, R. y Takaki, M. (2003) At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15, 130-163.
- Wax, R. H. (1971). *Doing Fieldwork: Warnings and Advice*. Chicago: University of Chicago Press.
- West, M. (1953). *A General List of English Words*. London: Longman.
- Wilkins, D. (1972). *Linguistics and Language Teaching*. London: Edward Arnold.
- Wingeyer, H. R. (2007). *Léxico disponible de Resistencia*. (Tesis doctoral inédita). Universidad de Alcalá de Henares.
- Wingeyer, H. R. y Angelina Oliveira, A. B. (2008). Resultados de la disponibilidad léxica de Corrientes y Resistencia: comparaciones desde un abordaje semántico de los términos referidos a 'el campo'. *Actas del IV Congreso Nacional sobre Problemáticas Sociales Contemporáneas*. Santa Fe, Universidad Nacional del Litoral.
- Wingeyer, H. R. y Angelina Oliveira, A. B. (2012). Identificación y análisis de las formas lingüísticas utilizadas para insultar en el español sincrónico de Corrientes. En E. M. Rojas Mayer (Coord.), *Léxico e interculturalidad: nuevas perspectivas* (pp. 1275-1278). San Miguel de Tucumán: Instituto de Investigaciones Lingüísticas y Literarias Hispanoamericanas.

- Wingeyer, H. R., Gusberti, J. V. y Trevisán, O. N. (2011). Análisis del tema: partes del cuerpo en la disponibilidad léxica de Resistencia y Corrientes. *Itinerarios Educativos*, 5(5), 64-74. Disponible en <http://bibliotecavirtual.unl.edu.ar/ojs/index.php/Itinerarios/article/viewFile/3937/5983>
- Wingeyer, H. y Gusberti, J. V. (2012). Presencia de unidades lingüísticas que remiten a mitos populares en la disponibilidad léxica de Resistencia y Corrientes. En E. M. Rojas Mayer (Coord.), *Léxico e interculturalidad: nuevas perspectivas* (pp. 1269-1274). San Miguel de Tucumán: Instituto de Investigaciones Lingüísticas y Literarias Hispanoamericanas.
- Wittgenstein, L. (1953). *Philosophical Investigations*. New York: McMillan.
- Wolfram, W. 1993. Ethical considerations in language awareness programs. *Issues in Applied Linguistics* 4(2), 225-255.
- Yager, K. (1998). Learning Spanish in Mexico: The effect of informal contact and students attitude on language gain. *Hispania*, 81(4), 898-913.
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Zahar, R., Cobb, T. y Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, 57(4), 541-572.
- Zohreh, K. y Saeedeh, E. (2011). The Effect of Bilinguality on L3 Breadth of Vocabulary Knowledge and Word Reading Skill. *Theory and Practice in Language Studies*, 1(8), 966-974.
- Zubarov, M. 2006. *Disponibilidad léxica en alemán e inglés como lengua extranjera*. (Tesis de magister inédita). Universidad de Concepción.

ANEXOS

Anexo I. Cuadro resumen de los objetivos, hipótesis y resultados de la tesis doctoral

Ámbito y pregunta de la investigación	Objetivos	Hipótesis	Resultados y conclusiones
Ámbito de la sociolingüística ¿Cuáles son los factores que inciden en el léxico virtual y en la identificación del léxico especializado?	Comparar el léxico disponible de este estudio con otros realizados para detectar semejanzas y diferencias por razones de lugar, tiempo, nivel de instrucción de los informantes, situación de la lengua (materna, extranjera o segunda lengua), entre otros.	El promedio de palabras de los estudiantes preuniversitarios malagueños será inferior al de los estudiantes de postgrado debido a la influencia del nivel de instrucción. El caudal léxico en lengua extranjera es inferior en lengua materna con independencia de la lengua de que se trate.	Media de los preuniversitarios 19,69 y de los estudiantes de postgrado 21,59 para los ocho centros de interés comunes. En español, los informantes han escrito un promedio de 191,91 palabras, en inglés 131,79 y en francés 147,71. Las comparaciones con otros estudios (Carcedo González, 2000; Ferreira, 2006; entre otros), muestran que en lengua materna el promedio de palabras siempre es superior a la lengua extranjera.
	Analizar las variables que condicionan el léxico disponible en español y en lengua extranjera del futuro profesorado para extraer datos que sean de interés en la enseñanza de idiomas.	Las estancias en el extranjero, el contacto con los medios de comunicación, la motivación por aprender lenguas y el mayor conocimiento de idiomas inciden positivamente en el caudal léxico.	En el caudal léxico en lengua extranjera las variables con mayor incidencia son: la especialidad del máster, el caudal léxico en español, la motivación del futuro profesorado y el uso de la lengua extranjera en diferentes contextos. En el caudal léxico en lengua española los resultados son más homogéneos ya que todos los informantes son estudiantes de postgrado y el nivel de estudios juega un papel fundamental. El resto de variables mayor influencia son: el nivel de instrucción, el caudal léxico en lengua extranjera y la especialidad del máster, (las especialidades varía con respecto a la lengua extranjera), la capacidad de identificación de tecnicismos.
	Conocer el dominio del léxico especializado en español y en lengua extranjera según los nueve centros de interés del diccionario de disponibilidad léxica y conocer cuáles son los factores que inciden en un mayor o menor conocimiento.	EL léxico disponible y la identificación de tecnicismos estará más influenciada por las diferentes variables en lengua extranjera que en español.	En lengua extranjera las variables con una incidencia positiva sobre el ICIT son el uso de la lengua en diferentes contextos, el ICIT en español, la elección del francés en la prueba, el ICoDD en lengua extranjera y la clase social. El conocimiento de otros idiomas guardan una relación significativa bilateral con el ICIT están presentes en los modelos multivariantes de forma indirecta, como puede ser a través del uso de la lengua y la especialidad del máster. En lengua española se produce la relación de influencia recíproca con el ICIT en lengua extranjera. El resto de variables destacadas

			son los hábitos de lectura, la exposición a programas culturales, la especialidad del máster y la clase social.
	Evaluar la relación que guarda el léxico virtual de los informantes y la identificación de tecnicismos.	Los resultados de la prueba de disponibilidad léxica y del léxico especializado guardan una estrecha relación de forma que a mayor caudal léxico mejor será el conocimiento de ese léxico especializado.	Los análisis estadísticos muestran una relación significativa a nivel inferior de 0,01 entre el ICoDD y el ICIT. Asimismo, la influencia de una variable sobre la otra queda patente en los análisis multivariantes.
Ámbito de la psicolingüística ¿Cómo se organiza el lexicón mental?	Evaluar las conexiones de palabras que se producen en el lexicón mental del informante a través del programa DispoGrafo (Echeverría <i>et al.</i> 2008) en español y en lengua extranjera.	Las palabras se organizan en redes. Cuando mayor sea la cohesión de un centro mejor definidas estarán las conexiones. La organización de las palabras en subgrupos está mejor definida en español que en lengua extranjera y en los centros más compactos con respecto a los más heterogéneos.	En los grafos se ha mostrado la organización de las palabras en subcampos léxicos, información que se complementa con los datos sobre las relaciones entre unidades léxicas. Efectivamente, los centros cohesionados dibujan unas redes más definidas y unas relaciones entre las palabras más fuertes.
	Realizar un análisis de los procesos de reentrada del léxico al núcleo o prototipo de cada centro de interés a través del índice de compatibilidad.	Los modelos de reentrada del léxico al núcleo varían en función del caudal léxico del informante y del tipo de centro de interés.	SE confirma la hipótesis. En los centros considerados más prototípicos se dan tres procesos de reentrada al núcleo: proceso de reentrada leve, moderado y pronunciado que corresponden a informantes con un bajo, medio y alto caudal léxico respectivamente. En cuanto a los centros relacionales, tan solo se localizan dos modelos de reentrada: modelo de reentrada moderado y pronunciado con desviación lineal.
	Esbozar una taxonomía de los centros de interés atendiendo a diferentes parámetros.	El índice de cohesión, el número de palabras y el número de vocablos viene determinado en cierto modo por el estímulo o centro de interés por lo que no todos tienen el mismo comportamiento.	Tras realizar análisis de los procesos de entrada al núcleo y de la prototipicidad de las palabras, constata que los resultados dependen del núcleo. Se propone una taxonomía de los centros de interés basado en un índice estructural que da cuenta de la homogeneidad del centro, de la organización de las palabras nucleares y de la fuerza de las conexiones entre los vocablos.. Cada centro se sitúa en un continuo que va desde centros prototípicos hasta centros relacionales.

Ámbito de la lingüística aplicada a la enseñanza de idiomas ¿Qué léxico enseñar?	Elaborar un diccionario de disponibilidad léxica en lengua española y en lengua extranjera (francés e inglés) sobre nueve centros de interés de gran relevancia en el ámbito educativo. Dentro de ese corpus, seleccionar las unidades léxicas que formarían parte de un vocabulario básico a partir de los listados del léxico disponible.	A través de la teoría de conjuntos difusos aplicada por Ávila Muñoz y Sánchez Sáez (2010, 2011) es posible seleccionar el léxico básico. Ese léxico corresponde al nivel de usuario básico establecido por el MCERL.	Se ha comprobado el alto grado de similitud entre el léxico nuclear en español, inglés y francés que ratifica que las palabras se sitúan en torno a un núcleo o prototipo. Se convierte así en una herramienta clave para aunar criterios en los estudios de disponibilidad léxica en la selección del vocabulario básico que se corresponde con los niveles A1 y A2 según la clasificación del Maro Común Europeo de Referencia para las Lenguas.
	Evaluar la corrección ortográfica de los listados de léxico disponible y conocer cuáles son los errores más comunes en los estudiantes para proponer medidas correctoras.	Los estudiantes de postgrado cometen considerablemente menos faltas de ortografía que los preuniversitarios. La mayoría de las disortografías se deben a una inadecuada acentuación de las palabras.	Los estudiantes de postgrado cometen una media de 0,28 palabras por informante y centro de interés en español y 0,73 en lengua extranjera que los preuniversitarios madrileños (Paredes García, 1999) cometen una media de 2,26 faltas de ortografía por informante y centro de interés y los preuniversitarios aragoneses 1,83 (Saura Rami, 2008). El 66,16 % de las disortografías corresponden a una inadecuada acentuación de las palabras
	Evaluar la idoneidad de los centros tradicionales en la creación de un léxico disponible.	Los centros de interés tradicionales no son suficientes para obtener el léxico disponible de un idioma.	Es bien sabido que existen centros poco productivos, como <i>Trabajos del campo y del Jardín e Iluminación, calefacción y medios de airear un recinto</i> , y que existen lagunas en los diccionarios del léxico disponible que dejan de lado aspectos básicos como la dimensión perceptiva y anímica del individuo. Se propone la utilización de las nociones específicas del Plan Curricular del Instituto Cervantes.
	Llevar a cabo una propuesta didáctica que se sirva de los resultados de los análisis realizados, sobre todo relacionados con la organización de las palabras en la mente.	Para elaborar una propuesta didáctica es necesario contar con análisis previos interdisciplinares.	Se ha realizado una propuesta pedagógica que se sirve de unos fundamentos teóricos sólidos y unos análisis que lo corroboran: análisis sobre la organización de las palabras en la mente, análisis sobre compatibilidad léxica y relaciones entre vocablos, cotejo de los centros de interés con las nociones específicas del Plan Curricular del Instituto Cervantes, entre otros aspectos.
	Indagar en el programa de estudios del Máster en Profesorado para conocer si la formación inicial de los docentes está orientada a alcanzar los objetivos del Plan de Fomento del Plurilingüismo.	No se exige una preparación previa ni metodológica ni lingüística que forme al futuro profesorado para impartir enseñanza bilingüe.	El profesorado carece de formación inicial en enseñanza bilingüe. El único requisito previo es acreditar un B1 mientras que en Andalucía se debe acreditar un B2 para impartir clases en lengua extranjera. La mayoría de los informantes desconoce conceptos básicos. Se propone la inclusión de una materia específica en el Máster en Profesorado.

Anexo II. Índice de gráficos, grafos y tablas

Gráfico 2.1. Distribución del léxico según categorías gramaticales en el estudio de Gougenheim <i>et al.</i> (1964)	56
Gráfico 2.2. Número de vocablos ordenados según el índice de compatibilidad en español. Centro de interés 09. <i>Ordenadores e internet</i>	113
Gráfico 2.3. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 145. Centro de interés 04. <i>The school: furniture and school material</i>	115
Gráfico 2.4. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 170. Centro de interés 04. <i>L'école: meubles et matériel scolaire</i>	115
Gráfico 2.5. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 032. Centro de interés 04. <i>Parties du corps humain</i>	116
Gráfico 2.6. Muestra de compatibilidad de vocablos. Patrón de centralizado. Informante 163. Centro de interés 01. <i>Partes del cuerpo humano</i>	116
Gráfico 2.7. Estadístico de resultados comparados. Fuente: Echeverría (2001: 94)	127
Gráfico 3.1. Niveles de población. Extraída de Rebagliato, Ruiz y Arranz (1996: 74)	162
Gráfico 4.1. Distribución de las palabras por centros de interés	215
Gráfico 4.2. Distribución porcentual del número de vocablos por centros de interés	216
Gráfico 4.3. Distribución porcentual de las palabras por centros de interés en lengua extranjera	219
Gráfico 4.4. Distribución porcentual de las palabras por centros de interés en inglés	220
Gráfico 4.5. Distribución porcentual de los vocablos por centros de interés en inglés	221
Gráfico 4.6. Distribución porcentual de las palabras por centros de interés en francés	222
Gráfico 4.7. Distribución porcentual de los vocablos por centros de interés en francés	223
Gráfico 4.8. Rangos según palabras, vocablos e índice de cohesión en español	225
Gráfico 4.9. Palabras, vocablos, densidad léxica e índice de cohesión en inglés	226
Gráfico 4.10. Comparación del promedio de palabras por centros de interés en español y en lengua extranjera	229
Gráfico 4.11. Comparación del promedio de palabras por centros de interés en inglés y francés	229
Gráfico 4.12. Distribución porcentual del tipo de errores en lengua española	255
Gráfico 4.13. Distribución porcentual de errores y palabras por centros de interés en español	256
Gráfico 4.14. Distribución porcentual de errores y palabras por centros de interés en lengua extranjera	258
Gráfico 4.15. Diagrama de cajas que recoge la distribución por sexos del número de palabras del centro denominado <i>La ropa</i>	268
Gráfico 5.1. Gráfico de tallos y hojas de la edad de los informantes	296
Gráfico 5.2. Diagrama de sectores de los estudios del informante	297
Gráfico 5.3. Diagrama de barras de los estudios de los padres	298
Gráfico 5.4. Diagrama de barras de la ocupación de los padres	300
Gráfico 5.5. Gráfico de sectores sobre los ingresos familiares	301
Gráfico 5.6. Diagrama de sectores del número de idiomas extranjeros que conoce el informante	306
Gráfico 5.7. Gráfico de sectores sobre la frecuencia en el uso de la lengua extranjera de la prueba en clase	307
Gráfico 5.8. Gráfico de sectores sobre el número de libros que el informante lee al año ...	308

Gráfico 5.9. Diagrama de sectores sobre la sensación del informante al impartir clases en lengua extranjera.....	309
Gráfico 5.10. Diagrama de cajas de la relación entre el ICoDD y el ICIT en LE.....	313
Gráfico 5.11. Diagrama de dispersión del ICoDD y el ICIT en lengua extranjera según el tipo de centro educativo.....	315
Gráfico 5.12. Diagrama de cajas que del ICoDD en español en relación con el ICIT por rangos	315
Gráfico 5.13. Diagrama de dispersión del ICoDD y el ICIT en español según el tipo de centro educativo	317
Gráfico 5.14. Diagrama de dispersión del ICIT en español y en lengua extranjera según el idioma de la prueba	318
Gráfico 5.15. Diagrama de dispersión del ICIT en lengua extranjera y en español según la movilidad del informante	320
Gráfico 5.16. Distribución de la media del índice corregido de identificación de tecnicismos con respecto al grado de estudios del padre.....	321
Gráfico 5.17. Distribución de la media del índice corregido de descentralización léxica en español con respecto al grado de estudios del padre	323
Gráfico 5.18. Diagrama de cajas de la distribución por áreas según el ICoDD en lengua extranjera.....	327
Gráfico 5.19. Promedio del ICoDD en lengua extranjera según la especialidad del máster	328
Gráfico 5.20. Diagrama de cajas del conocimiento de otras lenguas sobre el ICoDD en lengua extranjera.....	331
Gráfico 5.21. Diagrama de cajas del conocimiento de otras lenguas con respecto al ICoDD en lengua española.....	334
Gráfico 5.22. Distribución del ICoDD en español y en lengua extranjera con respecto al número de idiomas que conoce el informante.....	335
Gráfico 5.23. Distribución del ICIT en español y en lengua extranjera con respecto al número de idiomas que conoce el informante.....	336
Gráfico 5.24. Diagrama de cajas de la realización de la Erasmus con respecto al ICoDD en lengua extranjera.....	339
Gráfico 5.25. Diagrama de cajas de la realización de una estancia becada en el extranjero y el ICoDD en lengua extranjera.....	341
Gráfico 5.26. Diagrama de cajas de la realización de una estancia sin beca en el extranjero y el ICoDD en lengua extranjera	342
Gráfico 5.27. Promedio del ICIT en español según el número de libros que el informante lee al año	343
Gráfico 5.28. Distribución del ICoDD en lengua extranjera con respecto a la motivación para impartir clases en lengua extranjera	345
Gráfico 5.29. Gráfico de sedimentación del análisis factorial relacionado con la clase social	350
Gráfico 5.30. Pasos del modelo de regresión lineal del ICoDD en lengua extranjera.....	358
Gráfico 5.31. Pronóstico por observado del ICoDD en lengua extranjera.....	358
Gráfico 5.32. Histograma de residuos estudentizados del ICoDD en lengua extranjera.....	359
Gráfico 5.33. Medias estimadas del ICoDD en lengua extranjera según especialidades del máster	361
Gráfico 5.34. Medias estimadas del ICoDD en lengua extranjera y el ICoDD en español.	363
Gráfico 5.35. Medias estimadas entre el ICoDD en lengua extranjera y la sensación al impartir clases en lengua extranjera	364
Gráfico 5.36. Medias estimadas del ICoDD en lengua extranjera y el uso de la lengua extranjera en diferentes contextos	364
Gráfico 5.37. Pronóstico por observado del ICoDD en lengua extranjera	366

Gráfico 5.38. Histograma de residuos estudentizados del ICoDD en lengua extranjera	366
Gráfico 5.39. Medias estimadas del ICoDD en español y el ICoDD en lengua extranjera	368
Gráfico 5.40. Medias estimadas del ICoDD en español y la especialidad del máster	368
Gráfico 5.41. Medias estimadas del ICoDD y el ICIT en español	369
Gráfico 5.42. Pronóstico por observado del ICIT en lengua extranjera	370
Gráfico 5.43. Histograma de residuos estudentizados del ICIT en lengua extranjera	371
Gráfico 5.44. Medias estimadas del ICIT en lengua extranjera y española	372
Gráfico 5.45. Medias estimadas del ICIT y el ICoDD en lengua extranjera	373
Gráfico 5.46. Medias estimadas del ICIT en lengua extranjera y las asignaturas obligatorias en lengua extranjera en la carrera universitaria	373
Gráfico 5.47. Medias estimadas del ICIT en lengua extranjera y la lengua de la prueba	374
Gráfico 5.48. Medias estimadas del ICIT en lengua extranjera y la clase social	374
Gráfico 5.49. Pronóstico por observado del ICIT en español	375
Gráfico 5.50. Histograma de residuos estudentizados del ICIT en español	376
Gráfico 5.51. Medias estimadas del ICIT en español y la especialidad del máster	377
Gráfico 5.52. Medias estimadas del ICIT en español y en lengua extranjera	378
Gráfico 5.53. Medias estimadas del ICIT en español y las noticias y programas culturales en la radio	378
Gráfico 5.54. Medias estimadas del ICoDD en español y la especialidad del máster	379
Gráfico 5.55. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICoDD en lengua extranjera	384
Gráfico 5.56. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICoDD en español	387
Gráfico 5.57. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICIT en lengua extranjera	389
Gráfico 5.58. Diagrama del proceso causal de los efectos lingüísticos y extralingüísticos sobre el ICIT en lengua española	392
Gráfico 6.1. Vocablos ordenados según el índice de compatibilidad en español. Centro de interés 01. <i>Partes del cuerpo humano</i>	408
Gráfico 6.2. Niveles de prototipicidad del centro de interés <i>Comidas y bebidas</i>	410
Gráfico 6.3. Puerta de acceso a la red léxica: vocablos muy nucleares (Ávila Muñoz y Sánchez Sáez, 2010: 60)	413
Gráfico 6.4. Muestra de compatibilidad léxica de los 16 primeros vocablos. Informante 109. Centro de interés 01. <i>Partes del cuerpo humano</i>	413
Gráfico 6.5. Muestra de compatibilidad léxica de los 6 primeros vocablos. Patrón centralizado. Informante 157. Centro de interés 06. <i>Medios de transporte</i>	415
Gráfico 6.6. Muestra de compatibilidad léxica. Patrón de reentrada. Informante 170. Centro de interés 01. <i>Parties du corps humain</i>	415
Gráfico 6.7. Muestra de compatibilidad léxica. Patrón de desplazamiento y desviación. Informante 022. Centro de interés 04. <i>La escuela: muebles y materiales</i>	416
Gráfico 6.8. Muestra de compatibilidad léxica con los vocablos en posición 2, 3, 4 y 5. Patrón de desviación sistemática. Informante 119. Centro de interés 09. <i>Computer and Internet</i>	416
Gráfico 6.9. Muestra de compatibilidad léxica. Informante 062. Centro de interés 01. <i>Partes del cuerpo humano</i>	417
Gráfico 6.10. Escala de prototipicidad de los centros de interés	427
Gráfico 6.11. Muestra de compatibilidad léxica en español. Modelo de reentrada leve. Informante 060. Centro de interés 01. <i>Partes del cuerpo humano</i>	429
Gráfico 6.12. Muestra de compatibilidad léxica en inglés. Modelo de reentrada leve. Informante 145. Centro de interés 01. <i>Human body parts</i>	429

Gráfico 6.13. Muestra de compatibilidad léxica en francés. Modelo de reentrada leve. Informante 115. Centro de interés 01. <i>Parties du corps humain</i>	430
Gráfico 6.14. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 039. Centro de interés 01. <i>Partes del cuerpo humano</i>	431
Gráfico 6.15. Muestra de compatibilidad léxica en inglés. Modelo de reentrada moderado. Informante 161. Centro de interés 01. <i>Human body parts</i>	432
Gráfico 6.16. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 033. Centro de interés 01. <i>Parties du corps humain</i>	433
Gráfico 6.17. Muestra de compatibilidad léxica. Modelo de reentrada pronunciado. Informante 109. Centro de interés 01. <i>Partes del cuerpo humano</i>	434
Gráfico 6.18. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado. Informante 146. Centro de interés 01. <i>Human body parts</i>	436
Gráfico 6.19. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado. Informante 028. Centro de interés 01. <i>Parties du corps humain</i>	436
Gráfico 6.20. Muestra de compatibilidad léxica en español. Caso extremo. Informante 038. Centro de interés 01. <i>Partes del cuerpo humano</i>	437
Gráfico 6.21. Muestra de compatibilidad léxica en francés. Caso extremo. Informante 038. Centro de interés 01. <i>Parties du corps humain</i>	439
Gráfico 6.22. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 060. Centro de interés 04. <i>La escuela: muebles y materiales</i>	442
Gráfico 6.23. Muestra de compatibilidad léxica en inglés. Modelo de reentrada moderado. Informante 118. Centro de interés 04. <i>The school: furniture and school material</i>	442
Gráfico 6.24. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 170. Centro de interés 04. <i>L'école: meubles et matériel scolaire</i>	443
Gráfico 6.25. Muestra de compatibilidad léxica en español. Modelo de reentrada pronunciado con desviación lineal. Informante 060. Centro de interés 04. <i>La escuela: muebles y materiales</i>	444
Gráfico 6.26. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado con desviación lineal. Informante 146. Centro de interés 04. <i>The school: furniture and school material</i>	445
Gráfico 6.27. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado con desviación lineal. Informante 034. Centro de interés 04. <i>L'école: meubles et matériel scolaire</i>	446
Gráfico 6.28. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 060. Centro de interés 09. <i>Ordenadores e internet</i>	447
Gráfico 6.29. Muestra de compatibilidad léxica en español. Modelo de reentrada moderado. Informante 119. Centro de interés 09. <i>Computer and Internet</i>	448
Gráfico 6.30. Muestra de compatibilidad léxica en francés. Modelo de reentrada moderado. Informante 070. Centro de interés 09. <i>Ordinateurs et Internet</i>	448
Gráfico 6.31. Muestra de compatibilidad léxica en español. Modelo de reentrada pronunciado con desviación lineal. Informante 016. Centro de interés 09. <i>Ordenadores e internet</i>	449
Gráfico 6.32. Muestra de compatibilidad léxica en inglés. Modelo de reentrada pronunciado con desviación lineal. Informante 087. Centro de interés 09. <i>Computer and Internet</i>	451
Gráfico 6.33. Muestra de compatibilidad léxica en francés. Modelo de reentrada pronunciado con desviación lineal. Informante 034. Centro de interés 09. <i>Ordinateurs et Internet</i>	451
Gráfico 7.1. Resumen de las variables con mayor incidencia en el léxico en lengua extranjera	518
Gráfico 7.2. Resumen del modelo de regresión lineal del caudal léxico en español	519

Gráfico 7.3. Clasificación de los centros de interés de este estudio según el índice estructural	526
--	-----

Grafo 6.1. Centro de interés 01. <i>Partes del cuerpo humano</i>	460
Grafo 6.2. Centro de interés 01. <i>Human body parts</i>	461
Grafo 6.3. Centro de interés 01. <i>Parties du corps humain</i>	463
Grafo 6.4. Centro de interés 04. <i>La escuela: muebles y materiales</i>	464
Grafo 6.5. Centro de interés 04. <i>The school: furniture and school material</i>	465
Grafo 6.6. Centro de interés 04. <i>L'école: meubles et matériel</i>	466
Grafo 6.7. Centro de interés 09. <i>Ordenadores e internet</i>	467
Grafo 6.8. Centro de interés 09. <i>Ordenadores e internet</i> (ejemplo ampliado)	468
Grafo 6.9. Centro de interés 09. <i>Computer and Internet</i>	469
Grafo 6.10. Centro de interés 09. <i>Ordinateurs et internet</i>	471

Tabla 2.1. Ejemplo del listado de palabras clasificadas según orden de frecuencia. Extraído de Gougenheim <i>et al.</i> (1964: 69)	102
Tabla 2.2. Ejemplo del listado de vocablos ordenados según el número de menciones. Extraído de Gougenheim <i>et al.</i> (1964: 157)	103
Tabla 2.3. Listado de palabras con un mismo índice de disponibilidad pero diferente rango asociativo. Centro de interés 01. <i>Parties du corps humain</i> . Extraído de Njock (1978: 203)	106
Tabla 2.4. Ejemplo de los resultados del vocablo <i>hígado</i> en la investigación de Ávila Muñoz y Villena Ponsoda (2010) y el presente estudio	112
Tabla 2.5. Niveles de compatibilidad de los vocablos en español. Centro 09. <i>Ordenadores e internet</i>	114
Tabla 2.6. Informantes con mayor ICoAL e IAL en lengua española	118
Tabla 2.7. Ejemplos de respuestas en la prueba del léxico especializado en español y resultados del Pident e ICIT	123
Tabla 3.1. Plazas del Máster por especialidades y número de informantes	166
Tabla 3.2. Asignaturas del módulo general del Máster en Profesorado	171
Tabla 3.3. Áreas y especialidades del Máster en Profesorado	171
Tabla 3.4. Asignaturas del módulo específico del Máster en Profesorado	172
Tabla 3.5. Asignaturas del prácticum del Máster en Profesorado	172
Tabla 3.6. Asignaturas de libre configuración del Máster en Profesorado	172
Tabla 3.7. Centros de interés en los tres idiomas del estudio	179
Tabla 4.1. Número total de palabras y vocablos y sus respectivos promedios y rangos de productividad en español según el centro de interés	215
Tabla 4.2. Número total de palabras, promedio y rango de productividad en lengua extranjera	217
Tabla 4.3. Número de palabras y vocablos y sus respectivos promedios y rangos de productividad según el centro de interés	220
Tabla 4.4. Número de vocablos y rango por centro de interés en francés	222
Tabla 4.5. Densidad léxica e índice de cohesión en lengua española	225
Tabla 4.6. Palabras, Vocablos, densidad léxica e índice de cohesión en francés	227
Tabla 4.7. Promedio de palabras y rango por centros de interés en español, lengua extranjera, inglés y francés	230
Tabla 4.8. Resultados de los diez primeros informantes ordenados según el ICoAL en lengua extranjera	231
Tabla 4.9. Resultados de los diez primeros informantes ordenados según el ICoAL en lengua española	232

Tabla 4.10. Promedio de palabras y rango por centros de interés en español en tres estudios realizados en Málaga (Ávila Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010 y el estudio actual)	233
Tabla 4.11. Comparación de la media de palabras en inglés del estudio actual y de Dimitrijević (1969)	235
Tabla 4.12. Comparación de la media de palabras en inglés del estudio actual y de Ferreira (2006)	236
Tabla 4.13. Comparación del promedio de palabras en español y en el estudio de Carcedo (2000c)	236
Tabla 4.14. Comparación del promedio de palabras en español y en el estudio de Samper Hernández (2002)	237
Tabla 4.15. Comparación de los 20 vocablos más disponibles en español, inglés y francés	240
Tabla 4.16. Vocablos coincidentes entre los 20 más disponibles en español, francés e inglés	242
Tabla 4.17. Comparación de los veinte primeros vocablos en español de las investigaciones de Carcedo González (2000c), Samper Hernández (2002) y del estudio actual	244
Tabla 4.18. Tabla 4.18. Datos comparativos de las investigaciones de Carcedo González (2000c), Samper Hernández (2002) y del estudio actual	244
Tabla 4.19. Comparación de los veinte primeros vocablos en español de las investigaciones de Ávila Muñoz (2006), Ávila Muñoz y Villena Ponsoda (2010) y del estudio actual.	246
Tabla 4.20. Datos comparativos de las investigaciones de Ávila Muñoz (2006), Ávila Muñoz y Villena Ponsoda (2010) y del estudio actual	248
Tabla 4.21. Comparación de los veinte primeros vocablos en inglés de las investigación de Dimitrijević (1969) y del estudio actual	249
Tabla 4.22. Datos comparativos de la investigación de Dimitrijević (1969) y del estudio actual	250
Tabla 4.23. Comparación de los veinte primeros vocablos en inglés de las investigación de Ferreira (2006) y del estudio actual	251
Tabla 4.24. Datos comparativos de la investigación de Ferreira (2006) y del estudio actual	251
Tabla 4.25. Comparación de los veinte primeros vocablos en inglés de las investigación de Gougenheim <i>et al.</i> (1964) y del estudio actual	252
Tabla 4.26. Datos comparativos de la investigación de Gougenheim <i>et al.</i> (1964) y del estudio actual	252
Tabla 4.27. Clasificación de errores ortográficos en lengua española	254
Tabla 4.28. Palabras y vocablos incorrectos con respecto al total en español	257
Tabla 4.29. Palabras y vocablos incorrectos con respecto al total en inglés	259
Tabla 4.30. Palabras y vocablos incorrectos con respecto al total en francés	260
Tabla 4.31. Distribución de informantes según el número de faltas de ortografía en español y en lengua extranjera	260
Tabla 4.32. Tabla comparativa de los errores ortográficos del estudio de Paredes García (1999) y la muestra del presente estudio en español y en lengua extranjera	262
Tabla 4.33. Tabla comparativa de los errores ortográficos del estudio de Ávila Muñoz (2007) y la muestra del presente estudio en español y en lengua extranjera	263
Tabla 4.34. Tabla comparativa de los errores ortográficos del estudio de Saura Rami (2008) y la muestra del presente estudio en español y en lengua extranjera	263
Tabla 4.35. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en español	265
Tabla 4.36. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en inglés	266
Tabla 4.37. Cuadro resumen comparativo del número de vocablos aportados por hombres y mujeres en francés	266

Tabla 4.38. Media de vocablos en los diferentes idiomas según sexo	267
Tabla 4.39. Veinte primeros vocablos del centro 2, <i>La ropa</i> , según la variable sexo	269
Tabla 4.40. Vocablos aportados solo por hombres y mujeres en el centro 2, <i>La ropa</i>	270
Tabla 4.41. Veinte primeros vocablos del centro 3, <i>Comidas y bebidas</i> , según la variable sexo	271
Tabla 4.42. Vocablos aportados solo por hombres y mujeres en el centro 3, <i>Comidas y bebidas</i>	272
Tabla 4.43. Veinte primeros vocablos del centro 4, <i>La escuela: muebles y materiales</i> , según la variable sexo	273
Tabla 4.44. Vocablos aportados solo por hombres y mujeres en el centro 4, <i>La escuela: muebles y materiales</i>	274
Tabla 5.1. Análisis exploratorios sobre el ICoDD en lengua extranjera y en lengua española	293
Tabla 5.2. Comparativa del Pident y el ICIT en lengua extranjera por rangos	294
Tabla 5.3. Comparativa del Pident y el ICIT en español por rangos.....	295
Tabla 5.4. Análisis exploratorios de la edad de los informantes	296
Tabla 5.5. Tabla que muestra los diferentes estudios universitarios divididos por áreas.....	305
Tabla 5.6. Principales variables lingüísticas.....	310
Tabla 5.7. Principales variables sociodemográficas y reticulares	310
Tabla 5.8. Principales variables relacionadas con el aprendizaje de una lengua extranjera ..	311
Tabla 5.9. Datos descriptivos de los estudios del padre con respecto al índice corregido de identificación de tecnicismos en lengua extranjera.....	322
Tabla 5.10. Datos descriptivos de los estudios del padre con respecto al índice corregido de identificación de descentralización léxica en lengua extranjera	324
Tabla 5.11. Datos exploratorios del área del máster con respecto al ICoDD en lengua extranjera.....	326
Tabla 5.12. Análisis exploratorio del ICoDD en lengua extranjera por especialidades	328
Tabla 5.13. Análisis de la varianza de las especialidades del máster.....	329
Tabla 5.14. Análisis exploratorios del conocimiento de lenguas extranjeras con respecto al ICoDD en lengua extranjera.....	331
Tabla 5.15. Distribución de frecuencias según el número de lenguas extranjeras que conoce el informante	332
Tabla 5.16. Análisis exploratorios del del conocimiento de lenguas extranjeras con respecto al ICoDD en español	334
Tabla 5.17. Análisis de ANOVA de un factor de la variable relativa al conocimiento de otros idiomas	337
Tabla 5.18. Análisis exploratorios de la realización de la Erasmus con respecto al ICoDD en lengua extranjera	339
Tabla 5.19. Datos exploratorios de la realización de una estancia becada en el extranjero y el ICoDD en lengua	340
Tabla 5.20. Datos exploratorios de la realización de una estancia sin beca en el extranjero y el ICoDD en lengua extranjera.....	341
Tabla 5.21. Datos descriptivos de la motivación para impartir clases en lengua extranjera..	345
Tabla 5.22. Prueba de subconjuntos homogéneos que relaciona la sensación al impartir clases en lengua extranjera y el ICoDD en lengua extranjera	346
Tabla 5.23. KMO y prueba de Barlett del análisis factorial de la clase social.....	349
Tabla 5.24. Comunalidades del análisis de componentes principales de la clase social.....	349
Tabla 5.25. Prueba de la varianza total explicada del análisis de componentes principales de la clase social.....	350

Tabla 5.26. Prueba de la varianza total explicada del análisis de componentes principales de la motivación para impartir clases bilingües.....	352
Tabla 5.27. Resumen de la matriz de componentes en espacio rotado de la motivación para impartir clases bilingües.....	352
Tabla 5.28. Comunalidades del análisis de componentes principales del uso de la lengua...	354
Tabla 5.29. Comunalidades del análisis de componentes principales del uso de la lengua..	354
Tabla 5.30. Prueba de la varianza total explicada del análisis de componentes principales del uso de la lengua.....	355
Tabla 5.31. Datos estadísticos sobre el modelo de regresión lineal del caudal léxico en lengua extranjera.....	360
Tabla 5.32. Prueba de subconjuntos homogéneos que relaciona la especialidad del máster y el ICoDD en lengua extranjera	362
Tabla 5.33. Datos estadísticos sobre el modelo de regresión lineal del caudal léxico en español	367
Tabla 5.34. Datos estadísticos sobre el modelo de regresión lineal del léxico especializado en lengua extranjera	372
Tabla 5.35. Datos estadísticos sobre el modelo de regresión lineal del léxico especializado en español	376
Tabla 5.36. Efectos sobre las variables intermedias en el modelo causal del ICoDD en lengua extranjera	383
Tabla 5.37. Efectos en el modelo causal sobre la variable dependiente ICoDD en lengua extranjera.....	384
Tabla 5.38. Efectos sobre las variables intermedias en el modelo causal del ICoDD en lengua extranjera	386
Tabla 5.39. Efectos en el modelo causal sobre la variable dependiente ICoDD en lengua española.....	386
Tabla 5.40. Efectos sobre las variables intermedias en el modelo causal del ICIT en lengua extranjera.....	389
Tabla 5.41. Efectos en el modelo causal sobre la variable dependiente ICIT en lengua extranjera.....	389
Tabla 5.42. Efectos sobre las variables intermedias en el modelo causal del ICIT en lengua española.....	391
Tabla 5.43. Efectos en el modelo causal sobre la variable dependiente ICIT en lengua española.....	391
Tabla 6.1. Niveles de compatibilidad de los vocablos en español. Centro 01. <i>Partes del cuerpo humano</i>	411
Tabla 6.2. Niveles de compatibilidad de los vocablos en inglés. Centro 01. <i>Human body parts</i>	411
Tabla 6.3. Niveles de compatibilidad de los vocablos en francés. Centro 01. <i>Parties du corps humain</i>	411
Tabla 6.4. Clasificación de los centros según el número de vocablos (Gougenheim et al., 1964: 189-190).....	418
Tabla 6.5. Clasificación de los centros de interés del presente estudio según la estructura de organización interna y sus niveles de inclusividad.....	419
Tabla 6.6. Indicadores de la estructura de los centros de interés en español.....	423
Tabla 6.7. Indicadores de la estructura de los centros de interés en inglés	425
Tabla 6.8. Indicadores de la estructura de los centros de interés en francés.....	426
Tabla 6.9. Propuesta de clasificación de los centros de interés	427
Tabla 6.10. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica. Informante 109. Centro de interés 01. <i>Partes del cuerpo humano</i>	435

Tabla 6.11. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica en español. Informante 038. Centro de interés 01. <i>Partes del cuerpo humano</i>	438
Tabla 6.12. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica en francés. Informante 038. Centro de interés 01. <i>Parties du corps humain</i>	440
Tabla 6.13. Vocablos, índice de compatibilidad y posición para la muestra de compatibilidad léxica. Informante 016. Centro de interés 01. <i>Partes del cuerpo humano</i>	450
Tabla 6.14. Resultados comparativos en español en relación con palabras, vocablos y conexiones	455
Tabla 6.15. Resultados comparativos en inglés en relación con palabras, vocablos y conexiones	455
Tabla 6.16. Resultados comparativos en francés en relación con palabras, vocablos y conexiones	455
Tabla 6.17. Cliques y <i>clusters</i> más frecuentes en lenguas española	457
Tabla 6.18. Cliques más frecuentes en español. Centro de interés 01. <i>Partes del cuerpo humano</i>	459
Tabla 6.19. Tabla 6.20. Cliques más frecuentes en inglés. Centro de interés 01. <i>Human body parts</i>	461
Tabla 6.21. Cliques más frecuentes en francés. Centro de interés 01. <i>Parties du corps humain</i> .	462
Tabla 6.22. Cliques más frecuentes. Centro de interés 04. <i>La escuela: muebles y materiales</i>	464
Tabla 6.23. Cliques más frecuentes inglés. Centro de interés 04. <i>The school: furniture and school material</i>	465
Tabla 6.24. Cliques más frecuentes en francés. Centro de interés 04. <i>L'école: meubles et matériel</i>	466
Tabla 6.25. Cliques más frecuentes en español. Centro de interés 09. <i>Ordenadores e internet</i> .	467
Tabla 6.26. Cliques más frecuentes en inglés. Centro de interés 09. <i>Computer and Internet</i>	469
Tabla 6.27. Cliques más frecuentes en francés. Centro de interés 09. <i>Ordinateurs et Internet</i> .	470
Tabla 6.28. Vocablos aportados por un único informante y su proporción en el centro de interés con respecto al total de vocablos en lengua española	474
Tabla 6.29. Vocablos en español según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición	477
Tabla 6.30. Vocablos en inglés según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición	477
Tabla 6.31. Vocablos en francés según el índice de disponibilidad, frecuencia acumulada y porcentaje de aparición	478
Tabla 6.32. Léxico nuclear en español del centro de interés 01. <i>Partes del cuerpo humano</i>	479
Tabla 6.33. Léxico nuclear en inglés del centro de interés 01. <i>Human body parts</i>	479
Tabla 6.34. Léxico nuclear en francés del centro de interés 01. <i>Parties du corps humain</i>	479
Tabla 6.35. Tabla comparativa de los centros de interés y los temas y subtemas del Consejo de Europa, en Bartol Hernández (2010: 95).....	484
Tabla 6.36. Relación entre las nociones específicas y los centros de interés tradicionales	489
Tabla 6.37. Léxico nuclear en español del centro de interés 03. <i>Comidas y bebidas</i>	493
Tabla 6.38. Léxico nuclear en inglés del centro de interés 03. <i>Food and drink</i>	493
Tabla 6.39. Léxico nuclear en francés del centro de interés 03. <i>La nourriture et les boissons</i> ...	494
Tabla 6.40. Nociones específicas del nivel A1 y A2 del PCIC. Apartado 5.2. Bebida	495
Tabla 6.41. Nociones específicas del nivel A1 y A2 del PCIC. Apartado 5.3. Alimentos	495
Tabla 6.42. Nivel de referencia del léxico nuclear según el PCIC.....	497
Tabla 6.43. Ejemplo de la búsqueda del vocablo agua en la base de datos	500

Anexo III. Documento de negociación

Investigación sobre la evaluación de la competencia léxica en lengua inicial y en lengua extranjera a través del vocabulario virtual del futuro profesorado que cursa el Máster en Secundaria

1. Equipo de investigación

- Doctoranda: Inmaculada Clotilde Santos Díaz.
- Directores de tesis: Juan Andrés Villena Ponsoda, Catedrático de Lingüística General de la Universidad de Málaga. Antonio Manuel Ávila Muñoz, Profesor Titular del Departamento de Lingüística General de la Universidad de Málaga.

2. Naturaleza del trabajo

Este trabajo centra su investigación en la formación en lenguas extranjeras que tiene un alumnado específico: aquel que se convertirá tras un año de Máster en Profesorado. Nos interesa conocer el perfil del alumnado, la especialidad que cursa (sobre todo si es de un área de lengua extranjera o no), su nivel de competencia en una lengua extranjera, sus intereses y posibilidades para impartir clases bilingües o exclusivamente en un idioma extranjero y su valoración respecto a su formación.

Asimismo, indagaremos en el vocabulario disponible en lengua extranjera del alumnado en aras de centrarnos en dos objetivos diferentes pero complementarios. El primero se basaría en demostrar si la enseñanza de idiomas extranjeros aporta un vocabulario adecuado para que el futuro profesorado pueda impartir clases en ese idioma en cuestión. El segundo objetivo, consistirá en localizar los factores sociales, individuales y educativos que determinan la capacidad léxica de los sujetos analizados.

En la actualidad, el bilingüismo cobra cada vez más importancia en la educación y en el mercado laboral, por eso consideramos aconsejable hacer una indagación sistemática en el enfoque y en los resultados del sistema educativo español y, más concretamente en el sistema andaluz y en sus políticas lingüísticas de cara a la buena implantación del Plan de Fomento del Plurilingüismo aprobado por la Junta de Andalucía en 2005.

3. Iniciativa de la experiencia

Nuestras inquietudes por la formación lingüística del profesorado y nuestro afán en busca de medidas de mejora, nos ha llevado a escoger el Máster en Profesorado como foco de nuestra investigación ya que para lograr implementar con éxito el Plan de Fomento de Plurilingüismo es necesario conseguir una buena formación del profesorado tanto inicial como permanente.

El Máster en Profesorado se ha implantado en la Universidad de Málaga por primera vez en el curso 2009/2010 siendo el sustituto del Curso de Aptitud Pedagógica. Por consiguiente, es una titulación de postgrado que aporta la formación pedagógica y didáctica que habilita para el ejercicio de las profesiones de Profesor/a de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas.

Este máster nos permite el análisis de la formación inicial en lengua extranjera y lengua materna del futuro profesorado de toda la provincia de Málaga durante el curso 2010/2011. A través del análisis de la formación lingüística del alumnado del Máster en Profesorado podremos encontrar los puntos fuertes y débiles de nuestra política bilingüe así como intentar encontrar posibles medidas de mejora.

4. Metodología de la investigación

Dadas las características de esta investigación, utilizaremos el método cuantitativo y el cualitativo para obtener una visión más amplia y exacta de la realidad. Por una parte, el modelo cuantitativo nos aporta conocimientos de forma imparcial ya que se recogen y analizan los datos a través de conceptos y variables. Por otra parte, el modelo naturalista o cualitativo nos permite valorar y analizar mejor algo tan complejo como la opinión del alumnado del máster.

Los instrumentos que utilizaremos para recabar datos serán sobre todo el uso de cuestionarios, la confección de un diario de campo y el análisis de documentos elaborados al efecto. Cabe destacar que en los cuestionarios los/as alumnos/as podrán expresar y dar su opinión libremente así como ayudarnos a establecer categorías mediante una respuesta numérica.

5. Temporalización del trabajo

La recogida de datos de este trabajo se llevará a cabo durante el curso escolar 2010/2011 y posteriormente se pasará a su análisis y comentarios.

6. Condiciones de la investigación

En primer lugar, es importante señalar que no se utilizará documento o material alguno que no sea público sin la autorización previa de los interesados/as. Por ello, solicitaremos la autorización de los implicados/as en la investigación y acordaremos las condiciones de acceso a los cuestionarios que se desean pasar. .

En segundo lugar, es imprescindible recordar que se mantendrá el anonimato de los informantes y la confidencialidad de los datos. Asimismo, nadie tendrá acceso privilegiado a los datos de la investigación, aunque por su carácter público garantice a todos a conocer en su momento sus resultados.

No existe ningún estamento o persona cuya valoración de la realidad sea la correcta o cierta respecto a la formación del profesorado. Por tanto, pretendemos conocer y profundizar en distintas realidades que nos lleven a reflexionar sobre medidas de mejora a esta formación inicial. Puesto que esta investigación pretende ser una ayuda y, en ningún caso una amenaza, todos los participantes podrán tener conocimiento del informe emitido y lo podrán negociar en aras de evitar la subjetividad y la arbitrariedad pero sin vetarlo.

Por último, intentaremos causar las menos molestias posibles para no interferir en la marcha ordinaria de los cursos y poder seguir investigando y aportando nuevos datos a un programa de estudios tan reciente como es el Máster en Profesorado.

7. Consideraciones finales

Es importante que las personas implicadas en la investigación tengan conocimiento del trabajo que se está realizando. Por ello, es imprescindible hacerle llegar este documento de negociación a los coordinadores del máster así como al profesorado del grupo y a los representantes del colectivo de estudiantes.

Fdo. El profesor/a del grupo

Málaga, a 20 de enero 2011

12) Estudios del padre y de la madre:

Padre	Madre	Estudios: nivel más alto alcanzado
		0. Sin estudios
		1. Enseñanza primaria incompleta
		2. Enseñanza primaria completa
		3. Bachillerato o formación profesional de grado medio
		4. Diplomado o formación profesional de grado superior
		5. Licenciado
		6. Máster o doctorado

Observaciones (indique detalles complementarios que precisen la opción elegida, como por ejemplo **el número de años de estudio en enseñanza reglada**):

Padre:

Madre:

13) Ocupación del padre. Indique y/o describa su ocupación actual. Si está parado o jubilado indique **la última**.

14) Ocupación de la madre. Ídem.

15) ¿Cuál es la renta mensual familiar aproximada?

1. Hasta 400 2. De 400 a 799 3. De 800 a 110
 4. De 1101 a 1800 5. De 1801 a 3000 6. Más de 3000

16) Las personas con las que trata son (marque **solo una opción**):

1. Contactos solo en el barrio	
2. Contactos solo en Málaga/en su pueblo	
3. Contactos en Málaga y en pueblos de Málaga	
4. Contactos en Málaga y en otras capitales y pueblos de Andalucía	
5. Contactos en Málaga y en otras ciudades de España	
6. Contactos solo en otras ciudades de España	
7. Contactos dentro y fuera de España	

17) Densidad. Marque la respuesta que corresponda en cada uno de los apartados siguientes:

	Sí	No
1. Las personas a las que trata con frecuencia se conocen entre sí		
2. Las personas a las que trata con frecuencia se tratan mucho entre sí		
3. ¿Tiene usted en cuenta lo que piensan esas personas de sus propias decisiones antes de tomarlas? (Estudiar una determinada carrera, comprar un coche, etc.)		
4. ¿Les consulta antes de tomar una decisión?		
5. En caso de necesidad, ¿podría reunirlos a todos para un acontecimiento importante?		
6. ¿Se reúne usted con todos ellos alguna vez?		
7. ¿Se reúne usted con todos periódicamente?		

18) Movilidad. Marque la respuesta que corresponda en cada uno de los apartados siguientes:

	Sí	No
1. Las costumbres en el extranjero son mejores que en España		
2. Le gusta mucho viajar		
3. Se trasladaría a vivir al extranjero por motivos profesionales o personales		
4. Le gustaría trasladarse a vivir al extranjero una temporada		
5. Le gustaría trasladarse a vivir al extranjero para siempre		

19) Lengua materna del informante:

1. Español 2. Otra (escriba cuál): 20) Conocimiento de otras lenguas **del informante** (Indique el nivel según el *Marco Común Europeo de Referencia para las Lenguas (MCER)* si lo conoce):

	MCER	1. Inicial	2. Medio	3. Avanzado
Lengua:				
Lengua:				
Lengua:				
Lengua:				

21) Lengua materna del padre:

1. Español 2. Otra (escriba cuál):

22) Lengua materna de la madre:

1. Español 2. Otra (escriba cuál):

23) Indique qué lenguas extranjeras hablan los siguientes miembros de su familia:

	Lengua	MCER	1. Inicial	2. Medio	3. Avanzado
Padre					
Madre					
Hermano/a					
Pareja					
Otro (Indique quién):					

Observaciones: En caso de que el español no sea la lengua materna de algún miembro, indíquelo.

24) ¿En qué destreza tiene usted más dificultad en lengua extranjera? Marque una o varias.

Comprensión oral Comprensión escrita Expresión oral Expresión escrita

25) Uso de las diferentes lenguas. Indique con una cruz la opción adecuada añadiendo el número de personas con las que habla según las ocasiones.

ES: Español**LET:** Lengua Extranjera en la que se realiza el Test de disponibilidad (inglés/francés)**LE2:** Segunda Lengua Extranjera. En caso afirmativo, indique cuál:

	Todos los días			Varias veces por semana			Una vez por semana			Con menos frecuencia			Nunca		
	ES	LET	LE2	ES	LET	LE2	ES	LET	LE2	ES	LET	LE2	ES	LET	LE2
En clase															
Con amigos															
Con la familia															
Otro:															
Otro:															

26) Número de libros que lee al año.

En español En la lengua del test En otra lengua extranjera

27) Marque con una cruz el número de horas semanales que dedica a las siguientes actividades:

ES: Español

LET: Lengua Extranjera en la que se realiza el Test de disponibilidad (inglés/francés)

LE2: Segunda Lengua Extranjera

	Ninguna			Menos de 2			Entre 2 y 5			Más de 5		
	ES	LET	LE2	ES	LET	LE2	ES	LET	LE2	ES	LET	LE2
Videojuegos												
Búsqueda de información en Internet												
Comunicación en Internet (Messenger, chat, redes sociales tales como tuenti o facebook)												
Jugar en Internet												
Ver películas												
Vídeos												
Vídeos por Internet												
Escuchar música												
Horas de estudio												
Ver programas de televisión de noticias y documentales												
Ver programas de noticias del corazón y similares												
Leer periódicos locales												
Leer periódicos nacionales												
Oír programas de radio que incluyen noticias y programas culturales												

Formación en lenguas extranjeras

28) ¿Es usted especialista en lenguas extranjeras? 0. Sí ☐ 1. No ☐

29) ¿A qué edad empezó a estudiar la primera lengua extranjera?

30) ¿Dónde empezó a estudiar la primera lengua extranjera?

Colegio ☐ Otros (¿Dónde?):

31) ¿Le ha parecido difícil acreditar un nivel B1 para poder matricularse en este Máster?

Sí ☐ No ☐ No sabe/No contesta ☐

32) ¿Ha disfrutado de alguna beca para estudiar fuera de España? Sí ☐ No ☐

33) ¿Ha estudiado fuera de España (excluir becas)? Sí ☐ No ☐

34) ¿En qué centro de primaria ha estudiado?

0. Público ☐ 1. Concertado ☐ 3. Privado ☐ 4. Otros

35) ¿En qué centro de secundaria ha estudiado?

0. Público ☐ 1. Concertado ☐ 3. Privado ☐ 4. Otros

36) ¿Ha realizado la Erasmus? Sí ☐ ¿Por qué? Porque quería aprobar muchas asignaturas ☐

Porque le gustan las lenguas extranjeras ☐

Otro:

No ☐ ¿Por qué? Porque no le interesaba ☐

Porque le convalidaban pocas asignaturas ☐

Por no poder sufragar los gastos ☐

Porque no se la concedieron ☐

Otro:

37) ¿Ha cursado alguna asignatura obligatoria en lengua extranjera en su carrera? Sí ☐ No ☐

38) ¿Le han ofertado alguna asignatura optativa o de libre configuración en lengua extranjera? Sí

☐ No ☐

39) En caso afirmativo, ¿ha cursado alguna? Sí ☐ ¿Por qué? 1. Eran asignaturas fáciles ☐

2. Por motivos laborales ☐

3. Le gustan los idiomas ☐

4. Otro:

No ☐ ¿Por qué? 1. Eran asignaturas difíciles ☐

2. Por falta de utilidad ☐

3. No le gustan los idiomas

4. Otro:

40) ¿Ha estudiado alguna lengua extranjera durante sus estudios universitarios al margen de la universidad? Sí ☐ No ☐

41) En caso afirmativo, indique dónde y el número de años.

	Número de años
Centro de Idiomas de la Universidad	<input type="text"/>
Escuela Oficial de Idiomas	<input type="text"/>
Academias privadas	<input type="text"/>
Clases particulares	<input type="text"/>
Aprendizaje autodidacta (¿Cuál?):	<input type="text"/>
Otro (¿Cuál?):	<input type="text"/>

Información y opinión relacionada con el Plan de Fomento de Plurilingüismo

42) Solo en el caso de haber escogido el Máster en Profesorado de alguna especialidad de lengua extranjera, ¿en qué media imparte su profesor clase en ese idioma?

0. Solo en español (0% LE) ☐ 1. De 0 hasta 20% de la clase en LE ☐ 2. Entre 20-50% ☐

3. Entre 50-70% ☐ 4. Más de un 70% ☐ 5. 100% en LE ☐

43) ¿Le parecen suficientes las horas en lenguas extranjeras en el Máster? Sí ☐ No ☐

44) Solo en el caso de haber escogido el Máster en Profesorado de alguna especialidad diferente a la de lenguas extranjeras, ¿echa usted en falta alguna asignatura en lengua extranjera? Sí ☐

No ☐

45) ¿Escogerá un centro bilingüe para dar las prácticas? Sí ☐ No ☐ No lo sabe ☐

46) ¿Le gustaría trabajar en un centro educativo en el extranjero? Sí ☐ No ☐ No lo sabe ☐

47) ¿Qué lengua o lenguas le gustaría aprender o seguir aprendiendo?

1. Ninguna ☐ 2. Lengua: 3. No sabe/No contesta ☐

48) Antes de empezar sus estudios universitarios, ¿se había planteado ser profesor/a?

Sí ☐ No ☐ Otros:

49) ¿Le gustaría dar clases en un centro bilingüe?

¿Por qué?
 Sí ☐ 1. Le gustan los idiomas ☐
 2. Cree que el bilingüismo mejorará la enseñanza ☐
 3. Otro:

¿Por qué?
 No ☐ 1. No le gustan los idiomas ☐
 2. Cree que es mejor enseñar su asignatura en español ☐
 3. Cree que no está bien formado ☐
 4. Otro:

50) ¿Qué sensación le provoca a usted cuando recrea la imagen de verse en un aula teniendo que explicar en inglés su asignatura?

0. Ansiedad 1. Inquietud 2. Actitud neutra 3. Tranquilidad cer

51) ¿Sabe en qué consiste el AICLE (Aprendizaje Integrado de Contenidos en Lenguas Extranjeras)? Sí No

52) Cree usted que recibiendo las lecciones de matemáticas en inglés, el nivel de conocimientos finales del alumnado:

0. Empeorará 1. Será igual 2. Mejorará 3. No sabe/No contesta

53) ¿Considera que está bien formado para impartir clase dentro de un proyecto de bilingüismo? Marque de 0 (no está bien formado) a 5 (está muy bien formado).

0. 1. 2. 3. 4. 5.

54) ¿Cuál cree que sería su principal dificultad si tuviera que dar clases bilingües?

1. La falta de formación en lenguas extranjeras
 2. La falta de metodología adecuada para impartir clase
 3. Ambas
 4. Otra (¿Cuál?):

55) ¿Desea realizar algún comentario?

Anexo V. Pruebas de identificación de léxico especializado: español, inglés y francés

Test de léxico especializado

Sitúe en cada recuadro el **Centro de Interés (1-9)** en el que cree que corresponde cada palabra. Por ejemplo, *tórax* = 1 (partes del cuerpo). Si no sabe la respuesta deje el recuadro en blanco.

Proctólogo		Calesa	
Berlina		Megalópolis	
Conurbación		Cricket	
Taxidermista		Compilador	
Escroto		Halterofilia	
Mostachón		Isquion	
Enhebrar		Glasear	
Estola		Plóter	
Retroproyector		Pantógrafo	

Centros de interés

1. Partes del cuerpo
2. La ropa
3. Comidas y bebidas
4. La escuela (muebles y materiales)
5. La ciudad
6. Medios de transporte
7. Juegos y distracciones
8. Profesiones y oficios
9. Ordenadores e internet

Technical words test

Put in each box the number of the correct field. For instance: *thorax* = 1 (human body parts). If you don't know the answer, don't write it.

Chiropodist

--

Calash

--

Yacht

--

Citadel

--

Cenotaph

--

Rounders

--

Dustman

--

Phishing

--

Jawbone

--

Petanque

--

Stew

--

Spleen

--

Basting

--

Cashew

--

Mitten

--

Compiler

--

Petractor

--

Pantograph

--

Fields

1. Human body parts
2. Clothes
3. Food and drink
4. The school (furniture and school material)
5. The city
6. Means of transport
7. Games and entertainments
8. Professions and jobs
9. Computer and Internet

Test de lexique spécialisé

Mettez à côté de chaque mot le numéro du **Centre d'Intérêt (1-9)** correspondant. Par exemple :
tórax = 1 (parties du corps humain). Si vous ne connaissez pas la réponse, ne répondez pas.

Écuyer	<input type="text"/>	Gabarre	<input type="text"/>
Topographe	<input type="text"/>	Nécropole	<input type="text"/>
Minaret	<input type="text"/>	Bilboquet	<input type="text"/>
Dériveur	<input type="text"/>	Émulateur	<input type="text"/>
Urètre	<input type="text"/>	Agrès	<input type="text"/>
Ailler	<input type="text"/>	Apophyse	<input type="text"/>
Fautil	<input type="text"/>	Flamber	<input type="text"/>
Étole	<input type="text"/>	Progiciel	<input type="text"/>
Équerre	<input type="text"/>	Cellophane	<input type="text"/>

Centres d'intérêt

1. Parties du corps humain
2. Les vêtements
3. La nourriture et les boissons
4. L'école (meuble et matériel)
4. La ville
6. Moyens de transport
7. Jeux et loisirs
8. Professions et métiers
9. Ordinateurs et Internet

Anexo VI. Índice de siglas y abreviaturas

En español

ACP	Análisis de componentes principales
DEM	Diccionario del Español de México
DIAP	Dibujo, Imagen y Artes Plásticas
EEES	Espacio Europeo de Educación Superior
FOL	Formación y Orientación Laboral
IC	Índice de cohesión
ICe	Índice de centralidad
ICIT	Índice corregido de identificación de tecnicismos
ICo	Índice de conectividad
ICoDD	Índice corregido de descentralización léxica
LE	Lengua extranjera
MCER	Marco Común Europeo de Referencia para las Lenguas
n	Número de informantes
N0	Asociación arbitraria
N1	Palabra nuclear
N25	Palabra derivadas
N5	Asociación colectiva
PCIC	Plan Curricular del Instituto Cervantes
Pident	Proporción de identificación de tecnicismos
PPHDL	Proyecto Panhispánico de Disponibilidad Léxica
Ra	Rango asociativo
Rm	Rango medio
TIPI	Tecnología, Informática y Procesos Industriales

En inglés

Erasmus	European Region Action Scheme for the Mobility of University Students
FEV	Fuzzy Expected Value

SPSS Statistical Product and Service Solutions

En francés

CECRL Cadre européen commun de référence pour les langues

IC Indice de cohésion

Ice Indice de centralité

ICIMT Indice corrigé d'identification des mots techniques

ICo Indice de connectivité

ICoAL Indice corrigé d'amplitude lexicale

ICoDD Indice corrigé de décentralisation lexicale

Anexo VII. Clasificación de errores ortográficos en español, inglés y francés

Centro de interés	Tildes	Grafía c, z + e, i	j-g	s-x-c-k-qu	h	b-v	ll-y-ñ	otras
1. Partes del cuerpo humano	abdomen,apendice,biceps,celula,clavicula,corazon,cuadriceps, esofago, esternon, estomago, femur, gluteos, higado, indice, menton, michellin, musculo,oido, organo, órganos sexuales, ovulo, pancrea, pancreas, parpado, perone, pié, pomulo, pulmon, riñon, tendon, timpano, torax, traquea, triceps		vajina			bello, púvis, vaso	Menique, michellin	aruga, antemano, pulmo, pulmone, isquio
2. La ropa	algodon, boxer, calcetin, cardigan, cinturon, chaqueton, chandal, corcél, esmóquin, lenceria, pantalon, sueter, tacon	calzetines	trage			baquero, chuvasquero, vikini	maya	corro, fulá, jerseys, Legins, manga lasisa, pasmina, pijama, sonbrero, tennis
3. Comidas y bebidas	albondigas, ápio, atun, azucar, bebida isotonica, brocoli, calabacin, café, champan, champiñon, choped, choped, entrecó, fuagrás, infusion, jamon, judía, laudano, maiz, melocoton, melon, oregano, platano, salmon, sandia, sangria, te	calabazín	berengen a, potage		Ortaliza, zanaoria	aluvia, cerveza, bicarbonato		apsenta, chope, fondu, fuagrás, gaz, maíz, melogoton, mozzarella, pannini, pio nono, piza, sandwich, spagueti, wisqui, yogurt
4. La escuela: muebles y materiales	autobus, boligrafo, cafeteria, cañon, compas, consejeria, conserjeria, copisteria, estanteria, examenes, futbol, geografia, lampara, lapices, lapiz, liquido, portabló, portatil, posters, religion, secretaria, tipex		conserge, juguete	fiso		bate		clix, holl, mochilla, paragüero, pastilina, pizara, portamina, posit
5. La ciudad	almacen, arbol, arboles, autobus, cafeteria, callejon, carniceria, carril, cibercafé, claxón, comisaria, duplex, futbol, gran almacen, grua, jardin, lamparás, libreria, opera, pajaros,		agetreo	hogar de ansianos				Aparcamiento, casco antigüo, rotunda, théâtre, viaandante

	panadería, paseo marítimo, peluquería, policía, río, semáforo, tetería, tráfico, vehículo							
6. Medios de transporte	Autobús, avión, camión, Cercanías, helicóptero, gondola, helicóptero, máquina del tiempo, microbús, minibus, montopatin, omnibus, paracaídas, trailer, tranvía, trolebús, zepelín	cepelín		teleférico, trator			calluco	aladelta, cuá, monorail, pony, quack, scuter, tranvía, transatlántico
7. Juegos y distracciones	balónmano, colúmpio, cómic, dominó, elástico, espectáculo, ficción, fotografía, fútbol, música, ópera, playmóvil, pádel, parchis, película, policías, póker, póquer, poquer, tabu, televisión, triatlón, yo-yo						ralla	aljedrez, balet, balón mano, brica, fútbol, football, juegos de ordenador, kit, boxing, monopolio, paddle, póker, rayuela, tenis, video juego, voleibol
8. Profesiones y oficios	arquitecta técnica, arqueólogo, bailarín, broker, cardiólogo, chofer, conductor, conductor de autobús, coreógrafo, farmacéutico, filósofo, fotógrafo, gestor, ginecólogo, gruista, informático, intérprete, jeque árabe, mecánico, médico, músico, odontólogo, óptico, peón, policía, político, psicólogo, video			broquer, proseneta		avogado		afileador, armiero, barista, esteticista, franco tirador
9. Ordenadores e internet	batería, botón, cámara, dirección web, escáner, extraíble, inalámbrico, juegos de video, lápiz memoria, memoria RAM, módem, página web, portátil, raíz, ratón, router, video, vínculo					gravador, gravadora		Cooky, disquette, scanner, suscripción, webques

Centros de interés	Vocablos en inglés con las variantes erróneas
1. Human body parts	ankle (angle), arm (harm), belly button (belly botton), bone (bone), breast (breest, sens), bum (bump), cheek (chik), cranium (crain), ear (hear), elbow (helbow), eyebrow (eye brather, eyebrow), eyelash (eyelach, eyeleed, slash), feet (feeths, feets, fet), forehead (front-head, front head), hand (main), head (heat), heart (hart, hearth), kidney (cockney), knee (knie), leg (legg), lip (lib), mouth (mauth, mounth, mouse), neck (nek), nose (noice, noise, nouse), penis (pennis), shoulder (should), stomach (stommach), teeth (teeths, theef, thees, theet), thumb (dumb), thumbd), toe (too), tongue (longue, thongue, tonge, tong, tounge), tooth (thooth), vagina (vagen), wrist (whrist, wriste)
2. Clothes	accessories (accessories), anorak (anorall), blazer (blaiser, blasier), blouse (bluse), boot (boat, bote), bracelet (brazale), button (button), cotton (coton), dress (rove), dungarees (dungaree), gabaerdine (gabardin), glove (globe), hat (had), high heeled shoes (high, high heelled, high-heel shoes), jacket (chacket, chaquet, jacquet), jeans (jeens), knickers (knicker), leggings (leggin, legins), necklace (neckless), nickers (nicks), panties (panty), pants (pans), pullover (poulover, pulover), pyjamas (pigama, pijama, pijamas, piyama, pyjama, pyjamas), scarf (skarf), shoes (shous), skirt (skart), slippers (sleeper, sliper), sock (shock, shocks, snock, suck), sweater (sweter, swetter, swueter), swimming costume (swiming costume), T-shirt (t-short, t'shirt), thong (thong), trousers (trouser, trousseurs), underwear clothes (underwear clothe)
3. Food and drink	alcohol (alcohol), alcoholic drink (alcoholics drinks), apple (aple), aubergine (aubergene, auvergene, pin-eapple, pineapple), banana (banane), bean (been), beef (beaf, beff), beefsteak (beefsteck), beer (bear, bier, birra), bread (pain), broccoli (broccoly, brocoli), butter (but), champagne, cheese (chees, chese, chesse), chicken (chiken), chilli (chili), Chinese food (chinese food), chip (cheaps, ship), chocolate (chocolat), cinnamon (cinammon), cod (code), codfish (cod fish), coffee (caffé, coffé, coffe), coke (cola), cookie (cokie, cooky), crème brûlée (crème brullé), crêpe (crepe), dessert (desert), duck (dusk), fish and chips (fish in ships), hamburger (amburger, hamburg, hamburquer), jam (jamb), juice (jouice, juis), ketchup (ketchup), lemonade (limonade), lettuce (lettuce, letuce, luche), liquor (licor, licour), mandarine (mandarin), marmelade (mermelade), mayonnaise (mayonaise), meat (meet), melon (melo), mushroom (moussroom), mussel (musel), noodle (nuddle), nugget (nugget), omelette (omelette, ommelette), onion (oignon), pea (pee), peanut (penaut), pepper (peper), pineapple (ananas, peaneapple), pork chop (porkchop), potato (potatoe, potatos), raspberry (rawsberry), refresher (refresh), salad (salat), sandwich (sanwich), sausage (saussage, saysagge), seafood (sea food), snail (scargot), spaghetti (spagetti), strawberry (strawerry), tea (te, té, the), tomato (tomatoe, tomatos), watermelon (water-melon), whisky (wisky), wine (vine), yogurt (yofourt, yofurth, youghurt)
4. The school: furniture and school material	backpack (bagpack), bank (blanket), bicycle (bicicle), blackboard (blac table, blackboar), board (boar), chair (shair), classroom (class-room), connection (connexion), dictionary (diccionary), draw (drow), eraser (gome, gomme), gum eraser (eraser gum), gym (gim), gymnastic (gimnastic), mathematics (matematics), maths (mates, mats), notebook (note-book), pencilcase (pencase), physics (phisics), playground (background), professor (profesor), projector (projecter, proyector), punishment (punishment), pupil (puppil), racket (raquet), rubber (ruber), rucksack (rugsack), ruler (ruller), sharpener pencil (sharp pencil), sightsee (sighseen), stapler (stappler), student (eleve), task (tast), traffic (trafic), uncomfortable (uncomfortable),
5. The city	baker's (bakers), bike (byke), building (buildin), bus station (busstation), buses (busses), butcher's shop (butcher shop), bycicle (bicicle, bycicle), café (cofe), chemist's (chemists), citizen (citicien), city centre (centre city, the centre of the city), coffee (cooffee), cottage (coteage), facade (fachade), fireman (fire man), fortress (alcassar), fountain (fontain), garden (jardin), government (government), gypsy (gipsy), hypermarket (hipermarket), kiosk (kiosko), laundry (laundry), monument (munument), motorbike (motor-bike, motorbyke), motorcycle (motocycle, motorcicle), neighbourhood (heighbourhood, neigbourhood, neightbourhood, neirghbor), parliament (parlement), pavement (paviment), pedestrian crossing (pedestrian crosing), penthouse (pent-house), pollution (polution), rabbish (rubbish), railway (ralway), roundabout (round about, rounabout), rush hour (rush our), sea (see), secondary school (secundary school), shopping centre (shop centre, shoping centre), sightseeing (sightseeing), skyscraper (scraper, skyscrapper, skycraper, skyscraper), square (squere), stall (stoll), street (stret), supermarket (supermarkt), swimming (swimming), theatre (teather,

	teathre, theather), town hall (halltown), traffic (trafic), traffic jam (traffic jumps), zebra crossing (cebra crossing)
6. Means of transport	airplane (airplain), bicycle (bycicle, bycycle), bike (byke), boat (boot), bridge (pont), bus (autocar), canoe (canoo), chariot (carriot), cruise (croissier, cruce), cycle (cicle), donkey (dounke), flight (fleight), globe (glove), helicopter (helicopture, helycopter), hitch-hiking (autostop), jeep (yeap), lorry (camion), motorcycle (motocycle, motor-bike, motorcicle, mororciclye), mountain bike (mounteinbike), on foot (by feet, by foot), paraglider (paraglid), plane (plain), roller skate (skate rolling), skaters (skate), tram (tramp), transsiberian (transiverian), trolley (troley), wagon (vagoon), water polo (watherpool), yatch (jackt, yacth, yat, yate)
7. Games and entertainments	backgammon (backgamon), baseball (beisball, veisball), basketball (basketball, basket ball), beach (beack), bicycle (bycicle), biking (byking), billiard (biliard), blind man's buff (blindy hen), bungee (bunjee), chess (chest), computer game (computergame), cricket (criket), cycling (cicling), doll (dall), dominoes (domino), flute (floute), football (futball, futbol), gymnastics (gimnastic), hockey (hokey, jokey), kayak (cayak), ladders (snakes and stairs), listen to music (listen music), ludo (parchis, parchis), motorcycling (motorcicling), paddle (paddel, padel, padle), paragliding (parapenting), parcheesi (parchesse, parhees), Play Station (Play stación), playground (play ground), poker (pocker), ride on horse (ride a horse), rol game (role game), Scrabble (Scrabble), sculpture (esculpture), sightseeing (sighseeing), ski (sky), skiing (sking, skiing), soccer (socer, soccer), sport (deport), swim (swin), swimming pool (swim-pool, swiming pool), taboo (tabu), tennis (tenis), theatre (theathre, teatre, theatrer), volleyball (voleiball, voley-ball, voley-bal), watch (wach), World of Warcraft (word of craft)
8. Professions and jobs	accountant (accountman, countable), actress (actris), analist (analistic), architect (architecture, arquitech, architect), assessor (assesor), assistant (dependant), au pair (oper), baby-sitter (baby-siter, babysiter), boxer (boxet), broadcaster (broadcasting), bus driver (driver bus), bussinessman (bussinesman), chauffeur (choffeur, chofer, chofeur), chemist (quemist), cleaner (cleanner), commercial traveler (comercial traveler), computer specialist (informatic), construction worker (construccion work), cook (cooker), designer (dessiner), draftsman (draftman), draughtsman (drawman, drawer), economist (economist), electrician (electricist), electronical engineer (elctronic engineer), enfironmental (enviromental), engineer (engeneer, ingenier, engineeir, engineering, enginer, ingenier, enginner, ingeneer, ingenier, ingineer), entrepreneur (entreprener), fighter (fight man), fireman (fire man), fisherman (fisher), football player (football playe), geographer (geografer), guitarist (guitarrist), hairdresser (hair dresser), headmistress (headmister), hero (heroe), housekeeper (house keeper), illustrator (ilustrator), journalist (periodist), kitchen assistant (kitchent assistant), lawyer (lawer, layer, lywer), lieutenant (leutenant), locksmith (locker smith), mechanic (mechanical, mechannic), media man (media), member of parliament (member of parlament), military man (militar man), musician (musican, musicien), paddle player (padel player), pharmacist (farmachist), physical engineer (fisical engineer), physicist (phishy, physic), plumber (plumer), policeman (police, police man), politician (politicial, polititian), presenter (presentator), priest (pryest), professor (profesor, proffesor, proffessor), programmer (programer), psychologist (psichologist, psicogy, psicology, psicologist), salesman (saleman), scientific (cientific), scientist (scientific), shop assistant (shop assistance, shopp assistant), soccer player (soocer player), solo performer (concertist), sportsman (deportist, sportif), sportswoman (sportwoman), steward (stewars), tailor (taylor), taxi driver (taxist), translator (traductor), waiter (waitter), witch (wich), writer (writter)
9. Computer and Internet	access (acces), application (aplication), blog (bloggs), connection (conection, connexion), delete (delate, deleite), design (desing), desktop (deskop), dictionnary (dictionary), disk (disq), diskette (disquette), documentary (documental), download (dowload), Firefox (firewox), folder (foulder), Google traslator (traductor google), hard disk (harddisk), hard drive (hard driver), installation (instalation), keyboard (kepboard), keypad (keypan), laptop (labtop, lapton), mouse (mousse), page (pagine), photography fotograf), plug-in (plugin), printer (printter), processing (proccesing), programme (programm), received (recived), scanner (scaner), screen (screin, skreen), shopping (shoping), social net (social red), software (software), spyware (spiware), switch on (swiff on), Twitter (twiter), webcam (wedcam), website (web site), wifi (wiffi), World Wide Web (word wide wide),

Centros de interés	Vocablos en francés con las variantes erróneas
1. Parties du corps humain	avant-bras (avants bras), bras (brace), cheveux (cheveux), cul (cullote), dent (dente), doigt (doight, doit, doigts), hanche (anche), main (maine), nez (né), oeil (oeils, oil), oreille (oreile), rein (réin), sein (seine), sourcil (soucil), tête (tete), thorax (tora, torax)
2. Les vêtements	blouse (bluse), bonnet (bonet), botte (bôte), chaussette (chausette, chousette), chaussure (choissires), chemisier (chemisié), cravate (crevatte), foulard (foular), gant (globe), imperméable (impermeable), mocassin (mocasin), pyjama (pijama), soutien-gorge (soutien gorje), vêtement (vetement)
3. La nourriture et les boissons	alcool (alcohol), baguette (baguete), boeuf (veuf), carotte (carote), choucroute (choucroutte), citronnade (citronade), croissant (croisant, croissan), eau gazeuse (eau gazeuze), foie gras (foi gras), frite (fritte), girolle (girole), jus (juce), légume (legume), marmelade (mermelade), mousse (mouse), poisson (poison), profiterole (profiterolle), rhum (rom), saucisson (saucisson), sel (sail), vin (vine), yaourt (yahourt)
4. L'école: meubles et matériel	bâton (batôn), élève (élève), équerre (equerre), fenêtre (fenetre), fraîche (fraiche), mathématiques (mathematique, mathématique), rétroprojecteur (retroprojecteur), télévision (télévision)
5. La ville	appartement (apartement), autoroute (autourute), bâtiment (batîment), boulangerie (boulangérie), cathédrale (catédrale, cathedrale), centre urbain (centre urbaine), cimetière (cimentier), cordonnier (cordonier), département (department), discothèque (discotheque), hôtel (hotel), mairie (meri), métro (metro), musée (musé), pâtisserie (pâtisserie, patisserie), pharmacie (pharmatie), piéton (pieton), place (plâce), poissonnerie (poisonerie), rivière (riviere), rond point (rond poin), théâtre (théâtre), trottoir (trottoire)
6. Moyens de transport	à pied (a pied), avion (avión), bateau (bâteau), chevaux (cheveux), éléphant (elefant), hélicoptère (hélicoptère), métro (metro), motocyclette (motocyclete), péniche (peniche), sousmarin (sous marin), vélo (velo)
7. Jeux et loisirs	alcool (alcohol), balon prisonnier (balon prisonier), bibliothèque (biblioteque), cinéma (cinema), danser (dancer), élastique (elastique), Monopoly (monopoli, Monopolis), paddle (padel), pétanque (petanque), tarot (tarôt), théâtre (théatre), vélo (velo), vidéo (video), voyage (vouyage)
8. Professions et métiers	actrice (actrisse), architecte (arquitecte), charpentier (chapentier), chauffeur (chofeur, choffeur), commercial (comercial), cordonnier (cordonier), cuisinier (cuissinier), étudiant (etudiant), footballeur (footboleur), infirmier (enfermier), ingénieur (ingénieur), maître (maitré), maroquinier (marroquinier), médecin (médecin, medecin, medecin), pâtissier (patissier), pêcheur (poisonier), professeur (profeseur, proffesseur), réceptionniste (réceptionniste), retraité (jubilé), styliste (stiliste), traumatologiste (traumatologue)
9. Ordinateurs et internet	câble (cable), chater (chatear), connexion (conexion), courrier électronique (courier electronique), dossier (dosier), écran (ecran), écran (screen), électromagnétique (electromagnetic), hardward (hardward), imprimeur (imprimeur), opérateur (opérateur), Photoshop (photo shop), réseau (reseau), scanner (scaner), Skype (skipe), software (software), souris (mouse, souri), touche (touch)

Anexo VIII. Selección de vocablos nucleares en español, inglés y francés

Centro	Nivel	Índice de compatibilidad	Proporción	Número de vocablos	Vocablos
1. Partes del cuerpo humano	1	0,88	1,64	4	cabeza, ojo, mano, pie
	2	0,82	4,51	11	dedo, brazo, nariz, boca, oreja, pierna, pelo
	3	0,69	4,92	12	uña
	4	0,51	6,97	17	cara, cuello, ceja, hombro, codo
	5	0,33	11,07	27	diente, espalda, rodilla, pestaña, pecho, labio, lengua, tobillo, tronco, corazón
	6	0,18	17,62	43	muñeca, frente, estómago, mejilla, pulmón, cráneo, barriga, cadera, hueso, culo, antebrazo, cerebro, cintura, muslo, riñón, barbilla
2. La ropa	1	0,89	2,10	5	pantalón, camisa, falda, jersey, camiseta
	2	0,81	3,36	8	zapato, chaqueta, vestido
	3	0,67	4,2	10	calcetín, bufanda
	4	0,52	7,14	17	abrigo, sombrero, gorro, (pantalón) vaquero, media, rebeca, gorra
	5	0,34	11,35	27	braga, guante, calzoncillo, bota, sujetador, blusa, pañuelo, traje, corbata, cinturón
	6	0,17	17,23	41	sudadera, chaleco, chaquetón, pijama, tacón, chándal, zapatilla, tenis, tanga, bañador, sandalia, polo, chancla, camisón
3. Comidas y bebidas	1	0,84	0,19%	1	agua
	2	0,76	1,13%	6	cerveza, carne, pan, Coca-Cola, zumo
	3	0,62	2,07%	11	leche, pescado, patata, vino, café
	4	0,46	4,33	23	tomate, pasta, arroz, lechuga, paella, refresco, fruta, huevo, ensalada, lenteja, chocolate, verdura
	5	0,28	7,35%	39	manzana, tortilla, naranja, té, sopa, pollo, pizza, hamburguesa, potaje, filete, puchero, Fanta, espagueti, tortilla de patatas, bocadillo, plátano
	6	0,14	13,74%	73	macarrón, batido, zanahoria, jamón, galleta, ternera, limón, cebolla, güisqui, tarta, queso, fresa, legumbre, melón, atún, ron, pimienta, cereal, pera, aceite, chorizo, cerdo, sal, cocido, gazpacho, salchicha, calabaza, maíz, sandía, mantequilla, garbanzo, solomillo, azúcar, yogur
4. La escuela:	1	0,87	1,28	5	pizarra, mesa, silla, pupitre, lápiz

muebles y materiales	2	0,78	2,04	8	boli(grafo), tiza, libro
	3	0,64	2,81	11	profesor, goma (de borrar), ordenador
	4	0,46	4,08	16	libreta, borrador, regla, cuaderno, estuche
	5	0,27	7,4	29	clase, alumno, rotulador, escritorio, proyector, puerta, armario, mochila, sacapuntas, ventana, carpeta, folio, aula
	6	0,13	12,76	50	papel, escuadra, pizarra digital, compás, patio (de recreo), tarima, cartabón, estantería, biblioteca, maleta, lámpara, calculadora, escuela, perchero, agenda (escolar), recreo, papelera, gimnasio, cañón, pantalla, pegamento
5. La ciudad	1	0,86	0,62	3	calle, edificio, coche
	2	0,75	1,04	5	semáforo, casa
	3	0,61	1,86	9	plaza, carretera, parque, tienda
	4	0,44	3,73	18	avenida, centro comercial, gente, ayuntamiento, tráfico, acera, museo, (auto)bus, árbol
	5	0,27	7,25	35	restaurante, hospital, bar, cine, monumento, colegio, piso, farola, universidad, banco, teatro, metro, iglesia, moto(cicleta), aeropuerto, rotonda, barrio
	6	0,14	13,87	67	supermercado, panadería, ruido, paso de peatones, puerto, taxi, jardín, peatón, centro, estación, fuente, rascacielos, catedral, puente, escuela, atasco, persona, instituto, contaminación, playa, bici(cleta), tren, discoteca, estación de tren, bulevar, policía, estrés, polución, Correos, paseo marítimo, oficina, aparcamiento
6. Medios de transporte	1	0,9	3,16	6	coche, (auto)bus, avión, tren, bici(cleta), moto(cicleta)
	2	0,81	3,68	7	barco
	3	0,7	5,26	10	metro, taxi, tranvía
	4	0,51	6,84	13	metro, taxi, tranvía
	5	0,3	9,47	18	caballo, furgoneta, yate, monopatín, autocar
	6	0,16	14,74	28	avioneta, burro, submarino, patinete, ciclomotor, crucero, barca, coche de caballos, globo, tráiler
7. Juegos y distracciones	1	0,84	0,42	2	parchís, fútbol
	2	0,75	1,06	5	cine, baloncesto, (juego de) cartas
	3	0,61	1,9	9	ajedrez, videojuego, oca, leer
	4	0,44	3,81	18	música, televisión, deporte, tenis, internet, Trivial, juego de mesa, escondite, Monopoly
	5	0,27	7,19	34	póquer, salir con amigos, teatro, bailar, pillapilla, pasear, damas, petanca, viajar, natación,

					pádel, lectura, ordenador, balonmano, comba, correr
	6	0,14	13,53	64	pelota, mus, escuchar música, Tetris, Cluedo, cantar, libro, senderismo, película, bolo, atletismo, concierto, salir, brisca, serie, dominó, billar, pintar, elástico, esquiar, Play Station, patinar, escribir, gallinita ciega, voleibol, juego de ordenador, Tabú, bicicleta, videoconsola, tocar un instrumento
8. Profesiones y oficios	1	0,84	0,42	2	profesor, médico
	2	0,74	0,63	3	maestro
	3	0,6	1,67	8	ingeniero, abogado, enfermero, policía, arquitecto
	4	0,43	3,35	16	panadero, camarero, fontanero, pintor, traductor, carpintero, albañil, cocinero
	5	0,27	6,7	32	taxista, bombero, músico, actor, doctor, carnicero, conductor, informático, peluquero, electricista, político, cartero, psicólogo, administrativo, artista, barrendero
	6	0,15	14,64	70	limpiador, juez, periodista, futbolista, intérprete, mecánico, investigador, cantante, dependiente, dentista, ama de casa, director, conserje, agricultor, docente, escritor, jardinero, químico, constructor, recepcionista, secretario, camionero, banquero, pescadero, basurero, diseñador, filólogo, funcionario, vendedor, piloto, presentador, frutero, estudiante, arqueólogo, historiador, militar, tendero, arquitecto técnico
9. Ordenadores e internet	1	0,85	0,54	3	ratón, pantalla, teclado
	2	0,75	0,89	5	portátil, página web
	3	0,6	1,61	9	ordenador, monitor, Facebook, correo (electrónico)
	4	0,41	2,68	15	torre, impresora, red social, disco duro, programa, archivo
	5	0,24	5,89	33	Tuenti, internet, software, cable, web, hardware, pendrive, Google, blog, chatear, buscador, Windows, escáner, USB, chat, PC, red, wifi
	6	0,13	12,32	69	router, altavoz, módem, CD(-ROM), memoria, descargar, virus, placa base, tecla, descarga, conexión, ordenador de sobremesa, campus virtual, carpeta, navegar, TIC, pc, (Microsoft) Word, sistema operativo, documento, Twitter, procesador de textos, contraseña, DVD, foro, e-mail, videojuego, ordenador personal, información, ventana, CPU, antivirus, batería, netbook, Wikipedia, servidor

Centro	Nivel	Índice de compatibilidad	Proporción	Número de vocablos	Vocablos
1. Human body parts	1	0,91	4,35	5	head, eye, arm, leg, hand
	2	0,84	6,09	7	nose, mouth
	3	0,74	8,7	10	finger, ear, hair
	4	0,58	11,3	13	shoulder, foot, neck
	5	0,43	18,26	21	face, feet, knee, teeth, nail, toe, heart, elbow
	6	0,23	22,61	26	lip, tongue, back, eyebrow, chest
2. Clothes	1	0,9	3,31	5	trouser, T-shirt, skirt, jeans, shirt
	2	0,82	3,97	6	shoe
	3	0,7	5,3	8	jacket, hat
	4	0,54	8,61	13	sock, scarf, dress, jumper, pullover
	5	0,36	12,58	19	glove, coat, shorts, sweater, boot, pants
	6	0,19	18,54	28	bra, suit, tie, underwear, blouse, jersey, cardigan, belt, panties
3. Food and drink	1	0,86	0,7	2	water, meat
	2	0,77	1,39	4	Coke, fish
	3	0,65	2,79	8	wine, potato, tomato, beer
	4	0,51	6,62	19	milk, orange, hamburger, apple, pizza, chip, bread, vegetable, egg, rice, juice
	5	0,34	11,15	32	coffee, banana, carrot, soup, chicken, cake, Coca-Cola, fruit, lemon, tea, salad, lettuce, onion
	6	0,18	17,42	50	spaghetti, sandwich, beef, cheese, chocolate, bean, pasta, watermelon, sugar, pineapple, omelette, steak, whisky, strawberry, salt, pepper, soda, butter
4. The school: furniture and school material	1	0,89	2,04	5	table, chair, pencil, blackboard, pen
	2	0,8	2,86	7	desk, book
	3	0,65	3,27	8	teacher
	4	0,47	4,9	12	rubber, computer, notebook, paper
	5	0,28	7,76	19	ruler, chalk, window, classroom, rule, pupil, door
	6	0,14	13,06	32	student, board, pencil case, eraser, bag, map, class, workbook, ballpoint pen, maths, gym, exam, dictionary

5. The city	1	0,86	0,88	3	street, car, building
	2	0,76	1,17	4	bus
	3	0,62	1,75	6	house, park
	4	0,43	3,21	11	shop, road, people, square, traffic light
	5	0,28	7,87	27	cinema, restaurant, museum, church, underground, school, tree, flat, hospital, monument, supermarket, light, city centre, motorbike, skyscraper, traffic
	6	0,16	15,74	54	neighbourhood, shopping centre, pub, bank, taxi, bar, train, theatre, university, pollution, town hall, airport, avenue, noise, traffic jam, post office, bus station, train station, garden, pavement, city hall, bus stop, place, office, parking, cathedral, bridge
6. Means of transport	1	0,89	2,42	4	car, bus, train, plane
	2	0,8	3,03	5	bike
	3	0,7	5,45	9	motorbike, taxi, bicycle, underground
	4	0,51	6,67	11	boat, ship
	5	0,32	10,3	17	airplane, metro, horse, tube, helicopter, on foot
	6	0,16	15,76	26	motorcycle, skate, tram, lorry, cab, subway, yacht, aeroplane, truck
7. Games and entertainments	1	0,87	0,95	3	football, videogame, basketball
	2	0,78	1,26	4	cinema
	3	0,64	2,52	8	tennis, chess, television, music
	4	0,44	3,79	12	sport, card, Play Station, theatre
	5	0,28	7,26	23	reading, ball, volleyball, hide-and-seek, computer game, dance, film, computer, poker, movie, swimming
	6	0,14	14,2	45	soccer, baseball, running, concert, book, Wii, jogging, paddle, travel, walking, Trivial, hockey, rugby, Internet, Monopoly, dominoes, singing, Taboo, Scrabble, play, golf, crossword
8. Professions and jobs	1	0,87	1,07	3	teacher, doctor, policeman
	2	0,75	1,07	3	-
	3	0,28	2,14	6	nurse, lawyer, architect
	4	0,46	3,93	11	engineer, professor, waiter, plumber, drive
	5	0,28	7,86	22	painter, fireman, actor, writer, singer, butcher, taxi driver, carpenter, cook, secretary, musician

	6	0,16	15,71	44	mechanic, bus driver, manager, artist, translator, football player, barman, shop assistant, baker, journalist, pilot, psychologist, dancer, electrician, hairdresser, judge, designer, student, babysitter, salesman, assistant, politician
9. Computer and Internet	1	0,85	0,22	1	mouse
	2	0,75	0,86	4	screen, keyboard, web
	3	0,6	1,72	8	Facebook, laptop, computer, Google
	4	0,42	3,22	15	website, email, chat, hardware, software, Twitter, web page
	5	0,26	6,44	30	Internet, Tuenti, social network, download, blog, program(me), Windows, World Wide Web, pendrive, pc, printer, hard disk, CD(-ROM), Wi-Fi, game
	6	0,12	11,59	54	Messenger, monitor, desktop, personal computer, net, upload, link, Wikipedia, site, click, virus, video, Hotmail, (Microsoft) Word, scanner, video game, webcam, file, wiki, Youtube, mail, wireless, search, floppy disk

Centro	Nivel	Índice de compatibilidad	Proporción	Número de vocablos	Vocablos
1. Parties du corps humain	1	0,89	2,56	2	tête, main
	2	0,85	6,41	5	oeil, pied, bras
	3	0,76	10,26	8	nez, oreille, jambe
	4	0,61	14,1	11	bouche, cou, yeux
	5	0,44	19,23	15	doigt, cheveu, cheville, genou
	6	0,28	28,21	22	ongle, dent, ventre, dos, coude, langue, épaule
2. Les vêtements	1	0,9	3,75	3	chaussure, pantalon, jupe
	2	0,83	5	4	robe
	3	0,76	10	8	veste, jean, chemise, manteau
	4	0,62	15	12	chaussette, pull(óver), t(ee)-shirt, chapeau
	5	0,45	20	16	gant, bonnet, imperméable, echarpe
	6	0,29	27,5	22	cravate, blouse, culotte, casquette, pyjama, foulard
3. La nourriture et les boissons	1	0,85	0,6	1	eau
	2	0,79	1,81	3	fromage, chocolat
	3	0,65	3,01	5	omelette, bière
	4	0,53	7,83	13	jambon, lait, croissant, poisson, vin, pâte, pain, frite
	5	0,4	15,66	26	boeuf, oignon, canard, pomme, viande, thé, choucroute, citronnade, dîner, fruit, pomme de terre, orange, Coca-Cola
	6	0,23	22,29	37	vin blanc, pizza, crêpe, café, confiture, champagne, beurre, oeuf, poulet, légume, salade
4. L'école: meubles et matériel	1	0,91	3,54	4	table, crayon, chaise, tableau
	2	0,82	4,42	5	stylo
	3	0,69	5,31	6	livre
	4	0,53	7,97	9	cahier, professeur, bureau
	5	0,35	12,39	14	école, fenêtre, école primaire, salle de professeurs, sac
	6	0,23	22,12	25	feutre, gomme, ordinateur, dossier, élève, règle, papier, télévision, porte, sac à dos, concierge
5. La ville	1	0,88	1,48	2	voiture, rue

	2	0,79	2,22	3	maison
	3	0,68	4,44	6	bâtiment, police, route
	4	0,52	7,41	10	mairie, cinéma, avenue, place
	5	0,41	15,56	21	hôtel de ville, fontaine, poste, magasin, taxi, feu, boulangerie, centre, centre urbain, autobus, vélo
	6	0,24	23,7	32	restaurant, métro, église, train, appartement, boucherie, musée, quartier, centre ville, moto, parc
6. Moyens de transport	1	0,93	9,43	5	voiture, bus, train, vélo, avion
	2	0,87	11,32	6	métro
	3	0,78	13,21	7	taxi
	4	0,64	16,98	9	bateau, TGV
	5	0,48	20,76	11	tram(way), moto
	6	0,3	30,19	16	cheval, autobus, camion, RER, TER
7. Jeux et loisirs	1	0,89	2,36	3	cinéma, football, basket(-ball
	2	0,79	2,36	3	-
	3	0,68	3,94	5	musique, jeu vidéo
	4	0,53	7,09	9	rugby, théâtre, Monopoly, Cluedo
	5	0,38	14,17	18	poker, jeu de l'oie, volley-ball, base-ball, musée, tennis, cap ou pas cap?, Pro Evolution Soccer, fête
	6	0,22	19,69	25	sport, natation, vélo, Play Station, danser, discothèque, promenade
8. Professions et métiers	1	0,86	0,84	1	professeur
	2	0,8	2,52	3	boulangier, pompier
	3	0,7	5,88	7	plombier, policier, architecte, bouche
	4	0,52	7,56	9	médecin, jardinier
	5	0,35	11,77	14	cuisinier, ingénieur, infirmier, maître, traducteur
	6	0,24	23,53	28	serveur, docteur, informaticien, charcutier, pâtissier, cordonnier, chef, chauffeur, vendeur, gendarme, dentiste, militaire, charpentier, concierge
9. Ordinateurs et internet	1	0,89	2,34	3	souris, écran, clavier
	2	0,81	3,13	4	page web

	3	0,68	4,69	6	Internet, ordinateur
	4	0,51	6,25	8	logiciel, portable
	5	0,34	11,72	15	hardware, téléchargement, disquette, Facebook, informatique, clé USB, chat
	6	0,23	21,88	28	CD, Google, site (web), câble, software, JClic, Wi-Fi, page, imprimante, réseau social, Hot Potatoes, chatter, touche

ANEXO IX.
DICCIONARIOS DE DISPONIBILIDAD
LÉXICA EN ESPAÑOL, INGLÉS Y
FRANCÉS

Índice

1. Diccionarios en español.....	667
1.1. Listado de palabras en español por orden alfabético	667
1.1.1. Partes del cuerpo	667
1.1.2. La ropa	673
1.1.3. Comidas y bebidas.....	678
1.1.4. La escuela (muebles y materiales)	690
1.1.5. La ciudad.....	699
1.1.6. Medios de transporte	710
1.1.7. Juegos y distracciones	714
1.1.8. Profesiones y oficios	725
1.1.9. Ordenadores e internet.....	736
1.2. Listado de palabras en español ordenado según el índice de disponibilidad	749
1.2.1. Partes del cuerpo	749
1.2.2. La ropa	755
1.2.3. Comidas y bebidas.....	760
1.2.4. La escuela (muebles y materiales)	772
1.2.5. La ciudad.....	781
1.2.6. Medios de transporte	792
1.2.7. Juegos y distracciones	796
1.2.8. Profesiones y oficios	807
1.2.9. Ordenadores e internet.....	818
2. Diccionarios en inglés	831
2.1. Listado de palabras en inglés por orden alfabético	831
2.1.1. Human body parts.....	831
2.1.2. Clothes	834
2.1.3. Food and drink	838
2.1.4. The school (furniture and school material)	844
2.1.5. The city	850
2.1.6. Means of transport.....	858
2.1.7. Games and entertainments	862
2.1.8. Professions and jobs	869
2.1.9. Computer and internet	875
2.2. Listado de palabras en inglés ordenado según el índice de disponibilidad	886
2.2.1. Human body parts.....	886
2.2.2. Clothes	889
2.2.3. Food and drink	893
2.2.4. The school (furniture and school material)	899
2.2.5. The city	905
2.2.6. Means of transport.....	913
2.2.7. Games and entertainments	917
2.2.8. Professions and jobs.....	924
2.2.9. Computer and internet	930
3. Diccionarios en francés	941
3.1. Listado de palabras en francés por orden alfabético	941

3.1.1.	Parties du corps humain	941
3.1.2.	Les vêtements.....	943
3.1.3.	La nourriture et les boissons.....	945
3.1.4.	L'école (meubles et matériel)	949
3.1.5.	La ville	952
3.1.6.	Moyens de transport	955
3.1.7.	Jeux et loisirs	957
3.1.8.	Professions et métiers	960
3.1.9.	Ordinateurs et Internet.....	963
3.2.	Listado de palabras en francés ordenado según el índice de disponibilidad	
	966	
3.2.1.	Parties du corps humain	966
3.2.2.	Les vêtements.....	968
3.2.3.	La nourriture et les boissons.....	970
3.2.4.	L'école (meubles et matériel)	974
3.2.5.	La ville	977
3.2.6.	Moyens de transport	980
3.2.7.	Jeux et loisirs	982
3.2.8.	Professions et métiers	985
3.2.9.	Ordinateurs et Internet.....	988

1. Diccionarios en español

1.1. Listado de palabras en español por orden alfabético

1.1.1. Partes del cuerpo

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	abdomen	0,03769	0,312	0,312	8,187
2	abdominal	0,00976	0,111	0,423	2,924
3	alma	0,00272	0,022	0,445	0,585
4	alveolo	0,00310	0,045	0,490	1,170
5	ano	0,01151	0,111	0,601	2,924
6	antebrazo	0,11029	0,802	1,403	21,053
7	anular	0,00155	0,022	1,425	0,585
8	aorta	0,01121	0,089	1,514	2,339
9	aparato digestivo	0,00813	0,067	1,581	1,754
10	aparato excretor	0,00572	0,045	1,625	1,170
11	aparato fonador	0,00180	0,022	1,648	0,585
12	aparato reproductor	0,01204	0,111	1,759	2,924
13	apéndice	0,00491	0,045	1,804	1,170
14	apófisis	0,00502	0,022	1,826	0,585
15	arruga	0,00749	0,067	1,893	1,754
16	arteria	0,04306	0,445	2,338	11,696
17	articulación	0,00426	0,067	2,405	1,754
18	aurícula	0,00076	0,022	2,427	0,585
19	axila	0,01867	0,134	2,561	3,509
20	barba	0,01504	0,111	2,672	2,924
21	barbilla	0,08991	0,757	3,429	19,883
22	barriga	0,13959	1,113	4,542	29,240
23	bazo	0,01716	0,178	4,721	4,678
24	bíceps	0,03371	0,312	5,032	8,187
25	bíceps femoral	0,00210	0,022	5,055	0,585
26	bigote	0,01907	0,111	5,166	2,924
27	bilis	0,00084	0,022	5,188	0,585
28	boca	0,50902	2,917	8,105	76,608
29	brazo	0,49081	2,717	10,822	71,345
30	bronquio	0,00260	0,045	10,866	1,170
31	cabellera	0,00430	0,022	10,888	0,585
32	cabello	0,02818	0,178	11,067	4,678
33	cabeza	0,65670	2,895	13,961	76,023
34	cachete	0,00210	0,022	13,984	0,585
35	cadera	0,14499	1,202	15,186	31,579
36	capilar	0,00619	0,045	15,230	1,170
37	cara	0,20682	1,136	16,366	29,825
38	cardenal	0,00258	0,022	16,388	0,585
39	caries	0,00419	0,045	16,433	1,170
40	casco	0,00369	0,022	16,455	0,585

41	cava	0,00190	0,022	16,477	0,585
42	cavidad bucal	0,00453	0,022	16,500	0,585
43	ceja	0,26571	1,781	18,281	46,784
44	célula	0,00140	0,022	18,303	0,585
45	celulitis	0,00233	0,022	18,326	0,585
46	cerebelo	0,01005	0,067	18,392	1,754
47	cerebro	0,08533	0,646	19,038	16,959
48	cerumen	0,00369	0,022	19,060	0,585
49	cervical	0,00347	0,045	19,105	1,170
50	cintura	0,10338	0,779	19,884	20,468
51	clavícula	0,03685	0,289	20,174	7,602
52	clítoris	0,00258	0,022	20,196	0,585
53	codo	0,28046	1,937	22,133	50,877
54	cogote	0,00430	0,022	22,155	0,585
55	colmillo	0,01415	0,111	22,267	2,924
56	colon	0,00221	0,022	22,289	0,585
57	columna (vertebral)	0,03959	0,379	22,668	9,942
58	condición física	0,00200	0,022	22,690	0,585
59	corazón	0,20047	1,737	24,427	45,614
60	córnea	0,00505	0,045	24,471	1,170
61	costado	0,00133	0,022	24,493	0,585
62	costilla	0,01774	0,200	24,694	5,263
63	coxis	0,00140	0,022	24,716	0,585
64	cráneo	0,05157	0,289	25,006	7,602
65	cristalino	0,00140	0,022	25,028	0,585
66	cuádriceps	0,02172	0,223	25,251	5,848
67	cúbito	0,01071	0,111	25,362	2,924
68	cuello	0,33162	2,160	27,522	56,725
69	cuerda vocal	0,00523	0,067	27,589	1,754
70	cuero cabelludo	0,01689	0,089	27,678	2,339
71	cuerpo	0,00528	0,022	27,700	0,585
72	culo	0,12697	1,113	28,813	29,240
73	cutícula	0,00258	0,022	28,835	0,585
74	dedo	0,56195	3,162	31,997	83,041
75	dedo corazón	0,00221	0,022	32,020	0,585
76	dedo del pie	0,00445	0,045	32,064	1,170
77	dedo índice	0,01248	0,111	32,175	2,924
78	dedo pulgar	0,03862	0,289	32,465	7,602
79	deltoides	0,00126	0,022	32,487	0,585
80	deporte	0,00210	0,022	32,509	0,585
81	diente	0,27781	2,004	34,513	52,632
82	diente canino	0,00351	0,022	34,536	0,585
83	diente de leche	0,00147	0,022	34,558	0,585
84	empeine	0,00317	0,022	34,580	0,585
85	encia	0,00544	0,045	34,625	1,170
86	entrecejo	0,01818	0,111	34,736	2,924
87	entrepierna	0,00542	0,045	34,781	1,170
88	escroto	0,00843	0,045	34,825	1,170
89	esófago	0,02708	0,289	35,115	7,602

90	espalda	0,27851	2,138	37,252	56,140
91	espina dorsal	0,00771	0,067	37,319	1,754
92	espinilla	0,00508	0,067	37,386	1,754
93	esqueleto	0,01195	0,111	37,497	2,924
94	esternocleidomastoideo	0,01420	0,156	37,653	4,094
95	esternón	0,02570	0,267	37,920	7,018
96	estómago	0,15087	1,291	39,212	33,918
97	extremidad	0,04358	0,223	39,434	5,848
98	falange	0,01516	0,134	39,568	3,509
99	faringe	0,02840	0,312	39,880	8,187
100	fémur	0,03000	0,267	40,147	7,018
101	flequillo	0,00369	0,022	40,169	0,585
102	fosa nasal	0,00676	0,067	40,236	1,754
103	frente	0,11345	0,802	41,038	21,053
104	ganglio	0,00258	0,022	41,060	0,585
105	garganta	0,03753	0,379	41,438	9,942
106	gemelo	0,04983	0,445	41,884	11,696
107	genital	0,00858	0,089	41,973	2,339
108	glándula	0,00172	0,022	41,995	0,585
109	globo ocular	0,00477	0,022	42,017	0,585
110	glóbulo	0,00722	0,045	42,062	1,170
111	glóbulo blanco	0,00180	0,022	42,084	0,585
112	glóbulo rojo	0,00491	0,045	42,129	1,170
113	glúteo	0,00960	0,089	42,218	2,339
114	grano	0,00155	0,022	42,240	0,585
115	hígado	0,09071	0,824	43,064	21,637
116	hipotálamo	0,00180	0,022	43,086	0,585
117	hombro	0,30344	2,115	45,202	55,556
118	hoyuelo	0,00172	0,022	45,224	0,585
119	hueso	0,09022	0,735	45,959	19,298
120	huevo	0,00245	0,022	45,981	0,585
121	húmero	0,00210	0,022	46,003	0,585
122	incisivo	0,00389	0,022	46,025	0,585
123	ingle	0,00936	0,067	46,092	1,754
124	intestino	0,05382	0,534	46,627	14,035
125	intestino delgado	0,00613	0,089	46,716	2,339
126	intestino grueso	0,00629	0,089	46,805	2,339
127	iris	0,01628	0,156	46,961	4,094
128	isquion	0,00317	0,022	46,983	0,585
129	isquiotibial	0,00588	0,067	47,050	1,754
130	jugo gástrico	0,00590	0,067	47,116	1,754
131	labio	0,15121	1,002	48,118	26,316
132	laringe	0,03418	0,334	48,452	8,772
133	lengua	0,18389	1,291	49,744	33,918
134	lesión	0,00272	0,022	49,766	0,585
135	ligamento	0,00414	0,045	49,811	1,170
136	linfa	0,00200	0,022	49,833	0,585
137	lóbulo	0,03167	0,245	50,078	6,433
138	lumbar	0,00351	0,022	50,100	0,585

139	lunar	0,00658	0,089	50,189	2,339
140	mandíbula	0,00964	0,089	50,278	2,339
141	mano	0,60631	3,073	53,351	80,702
142	masa muscular	0,00163	0,022	53,373	0,585
143	médula espinal	0,00408	0,045	53,418	1,170
144	médula ósea	0,00351	0,022	53,440	0,585
145	mejilla	0,12948	0,980	54,420	25,731
146	menisco	0,00293	0,045	54,464	1,170
147	mentón	0,03625	0,289	54,754	7,602
148	meñique	0,01501	0,111	54,865	2,924
149	michelín	0,01098	0,111	54,977	2,924
150	moflete	0,00848	0,089	55,066	2,339
151	mucosa	0,00155	0,022	55,088	0,585
152	muela	0,03239	0,223	55,311	5,848
153	muela del juicio	0,00140	0,022	55,333	0,585
154	muñeca	0,15826	1,202	56,535	31,579
155	músculo	0,07375	0,579	57,114	15,205
156	muslo	0,10519	0,824	57,938	21,637
157	nalga	0,02097	0,178	58,116	4,678
158	nariz	0,52369	3,051	61,167	80,117
159	nervio	0,00351	0,022	61,189	0,585
160	nuca	0,03239	0,267	61,456	7,018
161	nudillo	0,03464	0,312	61,768	8,187
162	nuez	0,01208	0,111	61,879	2,924
163	oído	0,05106	0,312	62,191	8,187
164	ojo	0,67264	3,340	65,531	87,719
165	ombligo	0,08082	0,646	66,177	16,959
166	omóplato	0,00454	0,045	66,221	1,170
167	oreja	0,50052	2,961	69,183	77,778
168	órgano	0,01102	0,111	69,294	2,924
169	órgano sexual	0,00351	0,022	69,316	0,585
170	orificio nasal	0,01360	0,067	69,383	1,754
171	ovario	0,00951	0,111	69,495	2,924
172	óvulo	0,00093	0,022	69,517	0,585
173	paladar	0,00351	0,022	69,539	0,585
174	paleta	0,00602	0,045	69,584	1,170
175	palma de la mano	0,02123	0,134	69,717	3,509
176	páncreas	0,03662	0,334	70,051	8,772
177	pantorrilla	0,06157	0,557	70,608	14,620
178	panza	0,00369	0,022	70,630	0,585
179	papada	0,00155	0,022	70,652	0,585
180	párpado	0,02800	0,267	70,920	7,018
181	pecho	0,28416	2,093	73,013	54,971
182	pectoral	0,00519	0,045	73,057	1,170
183	pelo	0,39946	2,316	75,373	60,819
184	pelvis	0,03814	0,356	75,729	9,357
185	pene	0,03668	0,289	76,019	7,602
186	perilla	0,00258	0,022	76,041	0,585
187	peroné	0,01046	0,089	76,130	2,339

188	pestaña	0,23095	1,648	77,778	43,275
189	pezón	0,01608	0,134	77,911	3,509
190	pie	0,51059	3,140	81,051	82,456
191	piel	0,03288	0,334	81,385	8,772
192	pierna	0,55972	3,340	84,725	87,719
193	planta del pie	0,00877	0,089	84,814	2,339
194	plexo	0,00345	0,045	84,859	1,170
195	pómulo	0,03549	0,289	85,148	7,602
196	pubis	0,00831	0,089	85,237	2,339
197	punto del pie	0,00301	0,022	85,259	0,585
198	pulmón	0,14613	1,291	86,551	33,918
199	puño	0,00258	0,022	86,573	0,585
200	pupila	0,01837	0,178	86,751	4,678
201	radio	0,01228	0,134	86,885	3,509
202	recto	0,00210	0,022	86,907	0,585
203	riñón	0,10946	1,047	87,954	27,485
204	rodilla	0,30191	2,338	90,292	61,404
205	rostro	0,00415	0,045	90,336	1,170
206	rótula	0,00108	0,022	90,358	0,585
207	saliva	0,00443	0,045	90,403	1,170
208	sangre	0,02667	0,289	90,692	7,602
209	seno	0,00453	0,022	90,715	0,585
210	sexo	0,00301	0,022	90,737	0,585
211	sien	0,01193	0,089	90,826	2,339
212	talón	0,02936	0,334	91,160	8,772
213	tendón	0,01033	0,067	91,227	1,754
214	testículo	0,01071	0,089	91,316	2,339
215	teta	0,02282	0,200	91,516	5,263
216	tibia	0,01163	0,111	91,628	2,924
217	tímpano	0,03563	0,245	91,873	6,433
218	tiroides	0,00445	0,045	91,917	1,170
219	tobillo	0,20883	1,759	93,676	46,199
220	tórax	0,04501	0,312	93,988	8,187
221	torso	0,04984	0,334	94,322	8,772
222	tráquea	0,01545	0,223	94,545	5,848
223	trasero	0,01996	0,200	94,745	5,263
224	tríceps	0,01994	0,178	94,923	4,678
225	tripa	0,00405	0,045	94,968	1,170
226	trompa de Eustaquio	0,00301	0,022	94,990	0,585
227	trompa de Falopio	0,00507	0,045	95,035	1,170
228	tronco	0,09034	0,445	95,480	11,696
229	tubo digestivo	0,00400	0,045	95,524	1,170
230	uña	0,39057	2,627	98,152	69,006
231	uretra	0,00477	0,045	98,196	1,170
232	útero	0,01041	0,134	98,330	3,509
233	úvula	0,00317	0,022	98,352	0,585
234	vagina	0,02792	0,223	98,575	5,848
235	vejiga	0,01433	0,178	98,753	4,678
236	vello	0,02487	0,223	98,976	5,848

237	vena	0,04956	0,512	99,488	13,450
238	ventrículo	0,00072	0,022	99,510	0,585
239	vesícula	0,00280	0,045	99,555	1,170
240	vientre	0,04519	0,334	99,889	8,772
241	viscera	0,00200	0,022	99,911	0,585
242	vulva	0,00789	0,045	99,955	1,170
243	yugular	0,00233	0,022	99,978	0,585
244	zona lumbar	0,00258	0,022	100,000	0,585

1.1.2. La ropa

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(pantalón) pirata	0,00797	0,080	0,080	1,754
2	(pantalón) vaquero	0,23517	1,650	1,730	36,257
3	abalorio	0,00133	0,027	1,756	0,585
4	abrigo	0,36013	2,608	4,364	57,310
5	accesorio	0,00200	0,027	4,391	0,585
6	albornoz	0,02265	0,239	4,630	5,263
7	algodón	0,00619	0,053	4,683	1,170
8	alianza	0,00907	0,053	4,737	1,170
9	alpargata	0,00505	0,053	4,790	1,170
10	americana	0,00808	0,080	4,870	1,754
11	anillo	0,02183	0,213	5,082	4,678
12	anorak	0,05993	0,506	5,588	11,111
13	anteojo	0,00477	0,027	5,615	0,585
14	antifaz	0,00155	0,027	5,641	0,585
15	arreglarse	0,00163	0,027	5,668	0,585
16	babucha	0,01161	0,080	5,748	1,754
17	bailarina	0,01555	0,160	5,907	3,509
18	bañador	0,11615	1,064	6,972	23,392
19	bastón	0,00301	0,027	6,998	0,585
20	bata	0,05385	0,506	7,504	11,111
21	bata de cola	0,00221	0,027	7,531	0,585
22	bermudas	0,03010	0,213	7,743	4,678
23	biquini	0,05188	0,532	8,276	11,696
24	bisutería	0,00172	0,027	8,302	0,585
25	blazer	0,00369	0,027	8,329	0,585
26	blusa	0,16950	1,118	9,447	24,561
27	blusón	0,01241	0,106	9,553	2,339
28	body	0,00556	0,027	9,580	0,585
29	boina	0,02955	0,239	9,819	5,263
30	bolsa	0,00093	0,027	9,846	0,585
31	bolsillo	0,00458	0,053	9,899	1,170
32	bolso	0,05500	0,532	10,431	11,696
33	bombín	0,00839	0,053	10,484	1,170
34	Borsalino	0,00333	0,027	10,511	0,585
35	bota	0,22251	1,969	12,480	43,275
36	bota alta	0,00652	0,053	12,533	1,170
37	bota de agua	0,00792	0,080	12,613	1,754
38	bota de media caña	0,00333	0,027	12,640	0,585
39	bota de montaña	0,00333	0,027	12,666	0,585
40	botín	0,03635	0,319	12,986	7,018
41	botón	0,01136	0,106	13,092	2,339
42	bóxer	0,01849	0,186	13,278	4,094
43	braga	0,27961	2,288	15,567	50,292
44	braga (del cuello)	0,00155	0,027	15,593	0,585
45	brazalete	0,00466	0,053	15,647	1,170

46	broche	0,00172	0,027	15,673	0,585
47	bufanda	0,38735	2,821	18,494	61,988
48	burka	0,00599	0,053	18,547	1,170
49	burkini	0,00389	0,027	18,574	0,585
50	calcetín	0,47308	3,566	22,139	78,363
51	calza	0,00258	0,027	22,166	0,585
52	calzado	0,00477	0,027	22,193	0,585
53	calzón	0,00978	0,053	22,246	1,170
54	calzoncillo	0,26147	2,182	24,428	47,953
55	camisa	0,61591	3,619	28,047	79,532
56	camisa de dormir	0,00870	0,080	28,127	1,754
57	camiseta	0,61688	3,619	31,746	79,532
58	camiseta de tirantes	0,00221	0,027	31,772	0,585
59	camiseta interior	0,02625	0,266	32,038	5,848
60	camisola	0,00430	0,027	32,065	0,585
61	camisón	0,07321	0,559	32,624	12,281
62	capa	0,00756	0,080	32,704	1,754
63	capucha	0,00384	0,053	32,757	1,170
64	cárdigan	0,01454	0,160	32,916	3,509
65	casaca	0,00147	0,027	32,943	0,585
66	cazadora	0,05236	0,399	33,342	8,772
67	chal	0,03768	0,373	33,715	8,187
68	chaleco	0,16244	1,330	35,045	29,240
69	chamarra	0,00286	0,027	35,072	0,585
70	chamarreta	0,01035	0,080	35,152	1,754
71	chancla	0,07167	0,665	35,817	14,620
72	chándal	0,13884	1,224	37,041	26,901
73	chaqué	0,00450	0,053	37,094	1,170
74	chaqueta	0,49474	3,246	40,341	71,345
75	chaquetón	0,15527	1,171	41,511	25,731
76	chilaba	0,00496	0,053	41,565	1,170
77	chubasquero	0,06810	0,585	42,150	12,865
78	cinturón	0,18405	1,543	43,693	33,918
79	colgante	0,00520	0,080	43,773	1,754
80	collar	0,03302	0,293	44,066	6,433
81	colorido	0,00528	0,027	44,093	0,585
82	combinación	0,00172	0,027	44,119	0,585
83	complemento	0,00301	0,027	44,146	0,585
84	conjunto	0,00317	0,027	44,172	0,585
85	corbata	0,19370	1,570	45,742	34,503
86	cordón	0,00514	0,053	45,796	1,170
87	corpiño	0,01421	0,133	45,929	2,924
88	correa	0,01935	0,186	46,115	4,094
89	corsé	0,00508	0,053	46,168	1,170
90	cremallera	0,00820	0,080	46,248	1,754
91	culote	0,00981	0,080	46,328	1,754
92	delantal	0,00768	0,080	46,408	1,754
93	deportivo	0,00190	0,027	46,434	0,585
94	diminuto	0,00502	0,027	46,461	0,585

95	disfraz	0,00289	0,053	46,514	1,170
96	ejecutivo	0,00133	0,027	46,541	0,585
97	enagua	0,00432	0,053	46,594	1,170
98	equipación	0,00233	0,027	46,621	0,585
99	esmoquin	0,02403	0,213	46,833	4,678
100	estilo casual	0,00172	0,027	46,860	0,585
101	estola	0,01412	0,080	46,940	1,754
102	faja	0,00310	0,053	46,993	1,170
103	fajín	0,00120	0,027	47,020	0,585
104	falda	0,58022	3,539	50,559	77,778
105	falda de tubo	0,00389	0,027	50,585	0,585
106	falda de vuelo	0,00369	0,027	50,612	0,585
107	falda pantalón	0,00653	0,053	50,665	1,170
108	felpa	0,00922	0,080	50,745	1,754
109	formal	0,00200	0,027	50,772	0,585
110	forro	0,00210	0,027	50,798	0,585
111	forro polar	0,04613	0,399	51,197	8,772
112	frac	0,00582	0,080	51,277	1,754
113	fular	0,04350	0,399	51,676	8,772
114	gabardina	0,05913	0,506	52,182	11,111
115	gafa	0,02669	0,293	52,475	6,433
116	gafa de sol	0,01110	0,080	52,555	1,754
117	gargantilla	0,00456	0,053	52,608	1,170
118	gemelo	0,01818	0,160	52,767	3,509
119	gomilla	0,00133	0,027	52,794	0,585
120	gorra	0,18242	1,384	54,178	30,409
121	gorro	0,23632	1,836	56,014	40,351
122	guante	0,25640	2,076	58,089	45,614
123	hebilla	0,00286	0,027	58,116	0,585
124	impermeable	0,02274	0,213	58,329	4,678
125	informal	0,00190	0,027	58,356	0,585
126	jeans	0,00453	0,027	58,382	0,585
127	jersey	0,56256	3,326	61,708	73,099
128	jersey de cuello alto	0,00245	0,027	61,735	0,585
129	joya	0,00317	0,027	61,762	0,585
130	lana	0,00272	0,027	61,788	0,585
131	leggings	0,02245	0,239	62,028	5,263
132	lencería	0,00409	0,027	62,054	0,585
133	leotardo	0,04435	0,399	62,453	8,772
134	levita	0,03178	0,239	62,693	5,263
135	liguero	0,03482	0,293	62,986	6,433
136	mall	0,01372	0,133	63,119	2,924
137	manga	0,00301	0,027	63,145	0,585
138	manga a la sisa	0,00502	0,027	63,172	0,585
139	manoletina	0,02422	0,239	63,411	5,263
140	manopla	0,01753	0,133	63,544	2,924
141	mantilla	0,00172	0,027	63,571	0,585
142	mantón	0,00245	0,027	63,598	0,585
143	mantón de Manila	0,00323	0,053	63,651	1,170

144	media	0,29144	2,342	65,993	51,462
145	media tupida	0,00200	0,027	66,019	0,585
146	minifalda	0,06550	0,612	66,631	13,450
147	mitón	0,00133	0,027	66,658	0,585
148	mocasín	0,01527	0,133	66,791	2,924
149	mochila	0,00253	0,053	66,844	1,170
150	moda	0,00180	0,027	66,871	0,585
151	mono	0,01900	0,160	67,030	3,509
152	muñequera	0,00147	0,027	67,057	0,585
153	niqui	0,00477	0,027	67,084	0,585
154	orejera	0,01541	0,133	67,217	2,924
155	pajarita	0,04704	0,452	67,669	9,942
156	palabra de honor	0,00409	0,027	67,696	0,585
157	pamela	0,03915	0,293	67,988	6,433
158	pana	0,00258	0,027	68,015	0,585
159	pantalón	0,80262	4,337	72,352	95,322
160	pantalón chino	0,00602	0,053	72,406	1,170
161	pantalón corto	0,03023	0,239	72,645	5,263
162	pantalón de pana	0,00773	0,053	72,698	1,170
163	pantalón de pinzas	0,01019	0,080	72,778	1,754
164	pantufila	0,00351	0,027	72,805	0,585
165	panty	0,02040	0,160	72,964	3,509
166	pañuelo	0,16440	1,384	74,348	30,409
167	pañuelo palestino	0,00133	0,027	74,375	0,585
168	pareo	0,00507	0,053	74,428	1,170
169	pasamontañas	0,00391	0,053	74,481	1,170
170	pashmina	0,00798	0,106	74,588	2,339
171	peluca	0,00245	0,027	74,614	0,585
172	pendiente	0,02448	0,239	74,854	5,263
173	percha	0,00351	0,027	74,880	0,585
174	peto	0,02958	0,293	75,173	6,433
175	picardía	0,00320	0,053	75,226	1,170
176	pichi	0,00317	0,027	75,253	0,585
177	pijama	0,14543	1,251	76,503	27,485
178	plumón	0,02791	0,213	76,716	4,678
179	polo	0,06084	0,426	77,142	9,357
180	poncho	0,00670	0,080	77,222	1,754
181	probador	0,00221	0,027	77,249	0,585
182	pulsera	0,02363	0,239	77,488	5,263
183	rebeca	0,21181	1,517	79,005	33,333
184	reloj	0,01103	0,106	79,111	2,339
185	reloj de pulsera	0,00272	0,027	79,138	0,585
186	ropa	0,00818	0,053	79,191	1,170
187	ropa blanca	0,00233	0,027	79,218	0,585
188	ropa de color	0,00221	0,027	79,244	0,585
189	ropa deportiva	0,00503	0,053	79,297	1,170
190	ropa interior	0,07016	0,585	79,883	12,865
191	sandalia	0,07576	0,665	80,548	14,620
192	seda	0,00652	0,053	80,601	1,170

193	short	0,06001	0,532	81,134	11,696
194	slip	0,01156	0,133	81,267	2,924
195	sobrecamisa	0,00163	0,027	81,293	0,585
196	sombrero	0,21252	1,543	82,837	33,918
197	sombrero chambergó	0,00317	0,027	82,863	0,585
198	sombrero de copa	0,00140	0,027	82,890	0,585
199	sostén	0,00221	0,027	82,916	0,585
200	sudadera	0,11847	0,958	83,874	21,053
201	suéter	0,02187	0,160	84,034	3,509
202	sujetador	0,23713	1,889	85,923	41,520
203	tacón	0,13973	1,304	87,227	28,655
204	talla grande	0,00585	0,027	87,254	0,585
205	talla pequeña	0,00556	0,027	87,280	0,585
206	tanga	0,11956	1,064	88,345	23,392
207	tejano	0,01061	0,053	88,398	1,170
208	tenis	0,11214	0,958	89,356	21,053
209	tienda	0,00233	0,027	89,383	0,585
210	tirante	0,00933	0,133	89,516	2,924
211	toalla	0,00404	0,053	89,569	1,170
212	tocado	0,01135	0,080	89,649	1,754
213	top	0,04848	0,426	90,075	9,357
214	torera	0,00453	0,027	90,101	0,585
215	traje	0,12210	0,958	91,059	21,053
216	traje de baño	0,00575	0,053	91,112	1,170
217	traje de chaqueta	0,03367	0,293	91,405	6,433
218	traje de noche	0,00180	0,027	91,432	0,585
219	traje de vestir	0,00190	0,027	91,458	0,585
220	tres cuartos	0,02650	0,186	91,644	4,094
221	triquini	0,00544	0,053	91,698	1,170
222	túnica	0,00722	0,080	91,778	1,754
223	turbante	0,00453	0,027	91,804	0,585
224	uniforme	0,01110	0,133	91,937	2,924
225	velo	0,00539	0,053	91,990	1,170
226	vestido	0,39388	2,555	94,545	56,140
227	vestido camisero	0,00286	0,027	94,572	0,585
228	vestido de flamenca	0,00210	0,027	94,598	0,585
229	vestido de noche	0,00301	0,027	94,625	0,585
230	vestido de novia	0,00814	0,053	94,678	1,170
231	vestido de tirantes	0,00200	0,027	94,705	0,585
232	Wonderbra	0,00301	0,027	94,731	0,585
233	zanco	0,00163	0,027	94,758	0,585
234	zapatilla	0,13020	1,064	95,822	23,392
235	zapatilla de casa	0,00333	0,027	95,849	0,585
236	zapatilla de deporte	0,06039	0,506	96,354	11,111
237	zapato	0,51552	3,619	99,973	79,532
238	zueco	0,00155	0,027	100,000	0,585

1.1.3. Comidas y bebidas

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	absenta	0,00235	0,022	0,022	0,585
2	aceite	0,07552	0,635	0,657	16,959
3	aceituna	0,03406	0,263	0,920	7,018
4	aceituna rellena	0,00285	0,022	0,942	0,585
5	acelga	0,01795	0,131	1,073	3,509
6	agua	0,44275	2,256	3,330	60,234
7	agua con gas	0,00399	0,022	3,352	0,585
8	agua mineral	0,00380	0,022	3,373	0,585
9	aguacate	0,02344	0,175	3,549	4,678
10	aguardiente	0,00989	0,044	3,593	1,170
11	ajeno	0,00224	0,022	3,614	0,585
12	ajo	0,06149	0,504	4,118	13,450
13	ajo blanco	0,01068	0,088	4,206	2,339
14	albaricoque	0,01356	0,110	4,315	2,924
15	albariño	0,00483	0,022	4,337	0,585
16	albóndiga	0,02453	0,175	4,513	4,678
17	alcachofa	0,03324	0,219	4,732	5,848
18	alcohol	0,02281	0,131	4,863	3,509
19	alioli	0,00168	0,022	4,885	0,585
20	alita de pollo	0,00214	0,022	4,907	0,585
21	almeja	0,01837	0,131	5,038	3,509
22	almendra	0,01708	0,131	5,170	3,509
23	almuerzo	0,00259	0,022	5,192	0,585
24	alubia	0,03143	0,197	5,389	5,263
25	anacardo	0,00204	0,022	5,411	0,585
26	anchoa	0,02327	0,175	5,586	4,678
27	anguila	0,00247	0,022	5,608	0,585
28	anís	0,00645	0,044	5,652	1,170
29	aperitivo	0,00126	0,022	5,674	0,585
30	apio	0,02171	0,175	5,849	4,678
31	Aquarius	0,00992	0,044	5,893	1,170
32	arándano	0,00214	0,022	5,915	0,585
33	arroz	0,24437	1,533	7,448	40,936
34	arroz a la cubana	0,00829	0,066	7,514	1,754
35	arroz con leche	0,01301	0,110	7,623	2,924
36	asado	0,00109	0,022	7,645	0,585
37	atún	0,08261	0,635	8,280	16,959
38	ave	0,00857	0,044	8,324	1,170
39	avellana	0,00531	0,022	8,346	0,585
40	avena	0,00718	0,044	8,390	1,170
41	azafrán	0,00627	0,044	8,434	1,170
42	azúcar	0,06724	0,460	8,894	12,281
43	azúcar moreno	0,00345	0,022	8,916	0,585
44	bacalao	0,02367	0,175	9,091	4,678
45	Bacardi	0,01674	0,066	9,157	1,754

46	baguette	0,00399	0,022	9,179	0,585
47	banana	0,00778	0,044	9,222	1,170
48	barra de pan	0,00345	0,022	9,244	0,585
49	batata	0,00752	0,066	9,310	1,754
50	batido	0,10591	0,613	9,923	16,374
51	bebida	0,00314	0,022	9,945	0,585
52	bebida alcohólica	0,00473	0,044	9,989	1,170
53	bebida isotónica	0,00970	0,044	10,033	1,170
54	beicon	0,01097	0,131	10,164	3,509
55	berberecho	0,00329	0,022	10,186	0,585
56	berenjena	0,04418	0,263	10,449	7,018
57	berenjena con miel	0,00214	0,022	10,471	0,585
58	bicarbonato	0,00556	0,044	10,515	1,170
59	bistec	0,00185	0,022	10,537	0,585
60	bizcocho	0,02502	0,197	10,734	5,263
61	boca de mar	0,00247	0,022	10,756	0,585
62	bocadillo	0,06277	0,394	11,150	10,526
63	bollo	0,00559	0,066	11,216	1,754
64	bombón	0,00247	0,022	11,238	0,585
65	boquerón	0,04061	0,263	11,501	7,018
66	boquerón en vinagre	0,00676	0,044	11,544	1,170
67	brandy	0,00483	0,022	11,566	0,585
68	brécol	0,00380	0,022	11,588	0,585
69	brioche	0,00778	0,044	11,632	1,170
70	brócoli	0,00627	0,044	11,676	1,170
71	brote de soja	0,00132	0,022	11,698	0,585
72	brownie	0,00401	0,044	11,742	1,170
73	buey	0,00941	0,066	11,807	1,754
74	burrito	0,00224	0,022	11,829	0,585
75	caballa	0,00628	0,044	11,873	1,170
76	cacahuete	0,00194	0,022	11,895	0,585
77	cacao	0,00478	0,044	11,939	1,170
78	Cacique	0,00989	0,044	11,982	1,170
79	café	0,20338	1,139	13,122	30,409
80	calabacín	0,04988	0,394	13,516	10,526
81	calabaza	0,04568	0,285	13,801	7,602
82	calamar	0,03181	0,219	14,020	5,848
83	calamarito	0,00957	0,066	14,085	1,754
84	caldo	0,01131	0,088	14,173	2,339
85	caldo Avecrem	0,00380	0,022	14,195	0,585
86	caldo gallego	0,00483	0,022	14,217	0,585
87	callos	0,00869	0,044	14,261	1,170
88	calzone	0,00247	0,022	14,283	0,585
89	camarón	0,00362	0,022	14,304	0,585
90	canapé	0,00120	0,022	14,326	0,585
91	canela	0,01242	0,110	14,436	2,924
92	canelón	0,00686	0,066	14,502	1,754
93	cangrejo	0,00224	0,022	14,524	0,585

94	caña	0,00168	0,022	14,545	0,585
95	caqui	0,00364	0,044	14,589	1,170
96	caramelo	0,00784	0,066	14,655	1,754
97	cardo	0,00399	0,022	14,677	0,585
98	carne	0,32667	1,731	16,407	46,199
99	carne a la plancha	0,00345	0,022	16,429	0,585
100	carne de cerdo	0,00531	0,022	16,451	0,585
101	carne mechada	0,00214	0,022	16,473	0,585
102	carpaccio	0,00272	0,022	16,495	0,585
103	castaña	0,00995	0,088	16,583	2,339
104	cava	0,01443	0,088	16,670	2,339
105	caviar	0,00272	0,022	16,692	0,585
106	cazuela	0,00557	0,022	16,714	0,585
107	cazuela de fideos	0,00299	0,022	16,736	0,585
108	cebada	0,00863	0,066	16,802	1,754
109	cebolla	0,12478	0,964	17,766	25,731
110	cebolleta	0,00315	0,044	17,809	1,170
111	cena	0,00272	0,022	17,831	0,585
112	cerdo	0,09519	0,635	18,467	16,959
113	cereal	0,09162	0,657	19,124	17,544
114	cereza	0,02852	0,219	19,343	5,848
115	cerveza	0,32009	1,731	21,073	46,199
116	chacina	0,00362	0,022	21,095	0,585
117	champán	0,02181	0,131	21,227	3,509
118	champiñón	0,01536	0,131	21,358	3,509
119	chanquete	0,01887	0,131	21,490	3,509
120	Chardonnay	0,00506	0,022	21,512	0,585
121	chicle	0,00899	0,044	21,555	1,170
122	chipirón	0,00380	0,022	21,577	0,585
123	chirimoya	0,01791	0,131	21,709	3,509
124	chistorra	0,00451	0,044	21,752	1,170
125	choco	0,00204	0,022	21,774	0,585
126	chocolate	0,14068	0,854	22,629	22,807
127	chóped	0,01966	0,153	22,782	4,094
128	chopo	0,00194	0,022	22,804	0,585
129	chorizo	0,06287	0,438	23,242	11,696
130	chuchería	0,00603	0,044	23,286	1,170
131	chuleta	0,04063	0,263	23,549	7,018
132	chupachús	0,00439	0,022	23,571	0,585
133	chupito	0,00399	0,022	23,593	0,585
134	churro	0,00204	0,022	23,614	0,585
135	ciervo	0,00177	0,022	23,636	0,585
136	cigala	0,00345	0,022	23,658	0,585
137	ciruela	0,00556	0,088	23,746	2,339
138	Coca-Cola	0,21271	1,051	24,797	28,070
139	cochinillo	0,00132	0,022	24,819	0,585
140	cocido	0,04249	0,219	25,038	5,848
141	cocinar	0,00259	0,022	25,060	0,585
142	coco	0,00449	0,044	25,104	1,170

143	cóctel	0,00943	0,044	25,148	1,170
144	col	0,04469	0,307	25,455	8,187
145	Cola Cao	0,02398	0,175	25,630	4,678
146	coliflor	0,05276	0,394	26,024	10,526
147	comida basura	0,00483	0,022	26,046	0,585
148	concha fina	0,00439	0,022	26,068	0,585
149	condimento	0,00126	0,022	26,090	0,585
150	conejo	0,01280	0,088	26,177	2,339
151	conejo a la cazadora	0,00235	0,022	26,199	0,585
152	conserva animal	0,00299	0,022	26,221	0,585
153	conserva vegetal	0,00314	0,022	26,243	0,585
154	coñac	0,00886	0,044	26,287	1,170
155	coquina	0,00399	0,022	26,309	0,585
156	cordero	0,05073	0,307	26,616	8,187
157	costilla	0,01745	0,110	26,725	2,924
158	crema	0,01094	0,088	26,813	2,339
159	crema pastelera	0,00259	0,022	26,835	0,585
160	crepe	0,01103	0,066	26,900	1,754
161	criadilla	0,00460	0,022	26,922	0,585
162	croqueta	0,01982	0,131	27,054	3,509
163	cruasán	0,01306	0,088	27,141	2,339
164	cuajada	0,00146	0,022	27,163	0,585
165	cubalibre	0,00399	0,022	27,185	0,585
166	cuchara	0,00068	0,022	27,207	0,585
167	cuchillo	0,00065	0,022	27,229	0,585
168	cuscús	0,00513	0,044	27,273	1,170
169	Danone	0,00314	0,022	27,295	0,585
170	desayuno	0,00816	0,044	27,338	1,170
171	dorada	0,00329	0,044	27,382	1,170
172	Doritos	0,00585	0,022	27,404	0,585
173	dulce	0,02509	0,219	27,623	5,848
174	Dyc	0,00557	0,022	27,645	0,585
175	emblanco	0,00532	0,044	27,689	1,170
176	embutido	0,02191	0,175	27,864	4,678
177	empanada	0,01352	0,088	27,952	2,339
178	empanadilla	0,00482	0,044	27,996	1,170
179	ensalada	0,19421	1,183	29,179	31,579
180	ensalada de arroz	0,00153	0,022	29,200	0,585
181	ensaladilla rusa	0,01114	0,110	29,310	2,924
182	entrante	0,00506	0,022	29,332	0,585
183	entrecot	0,01398	0,110	29,441	2,924
184	escarola	0,00399	0,022	29,463	0,585
185	espagueti	0,09201	0,504	29,967	13,450
186	espárrago	0,01946	0,110	30,077	2,924
187	especia	0,00962	0,088	30,164	2,339
188	espeto	0,00345	0,022	30,186	0,585
189	espinaca	0,02694	0,197	30,383	5,263
190	estofado	0,03752	0,219	30,602	5,848
191	fabada	0,02672	0,175	30,778	4,678

192	Fanta	0,06550	0,329	31,106	8,772
193	Fanta Limón	0,00483	0,022	31,128	0,585
194	Fanta Naranja	0,00506	0,022	31,150	0,585
195	fiambre	0,00374	0,044	31,194	1,170
196	fideo	0,03531	0,263	31,457	7,018
197	fideuá	0,00299	0,022	31,479	0,585
198	filete	0,11021	0,635	32,114	16,959
199	filete tártaro	0,00214	0,022	32,136	0,585
200	flamenquín	0,00856	0,066	32,202	1,754
201	flan	0,01373	0,110	32,311	2,924
202	flan de huevo	0,00272	0,022	32,333	0,585
203	foie gras	0,00996	0,088	32,421	2,339
204	fondue	0,01173	0,066	32,486	1,754
205	frambuesa	0,01653	0,131	32,618	3,509
206	fréjol	0,00362	0,022	32,640	0,585
207	fresa	0,11083	0,811	33,450	21,637
208	fruta	0,20306	1,161	34,611	30,994
209	fruto rojo	0,00235	0,022	34,633	0,585
210	fruto seco	0,01002	0,066	34,699	1,754
211	fusilli	0,00483	0,022	34,721	0,585
212	galleta	0,06106	0,416	35,137	11,111
213	gallina	0,00224	0,022	35,159	0,585
214	gallo	0,00224	0,022	35,181	0,585
215	gamba	0,03141	0,241	35,422	6,433
216	gamba rebozada	0,00299	0,022	35,444	0,585
217	garbanzo	0,04829	0,372	35,816	9,942
218	gaseosa	0,01894	0,131	35,947	3,509
219	gazpacho	0,05711	0,329	36,276	8,772
220	gazpachuelo	0,00956	0,044	36,320	1,170
221	gelatina	0,00204	0,022	36,342	0,585
222	ginebra	0,01267	0,088	36,429	2,339
223	gin-tonic	0,01651	0,088	36,517	2,339
224	gofre	0,00160	0,022	36,539	0,585
225	golosina	0,00371	0,044	36,583	1,170
226	gominola	0,00621	0,044	36,627	1,170
227	granada	0,00765	0,066	36,692	1,754
228	granizada	0,02167	0,131	36,824	3,509
229	guarnición	0,00259	0,022	36,846	0,585
230	guisante	0,05544	0,416	37,262	11,111
231	guiso	0,02047	0,088	37,349	2,339
232	güisqui	0,07139	0,416	37,766	11,111
233	gula	0,00235	0,022	37,788	0,585
234	haba	0,03826	0,263	38,050	7,018
235	habichuela	0,01895	0,153	38,204	4,094
236	hambre	0,00247	0,022	38,226	0,585
237	hamburguesa	0,09424	0,526	38,751	14,035
238	harina	0,02550	0,197	38,949	5,263
239	helado	0,05161	0,394	39,343	10,526
240	hielo	0,00160	0,022	39,365	0,585

241	hierbabuena	0,00407	0,044	39,409	1,170
242	hígado	0,00483	0,044	39,452	1,170
243	hojaldre	0,00285	0,022	39,474	0,585
244	horchata	0,00362	0,022	39,496	0,585
245	hortaliza	0,01545	0,131	39,628	3,509
246	huevo	0,17571	1,073	40,701	28,655
247	huevo cocido	0,00329	0,022	40,723	0,585
248	huevo de codorniz	0,00314	0,022	40,745	0,585
249	infusión	0,02953	0,219	40,964	5,848
250	ingrediente	0,00194	0,022	40,986	0,585
251	jabalí	0,00380	0,022	41,008	0,585
252	jamón	0,14388	1,117	42,125	29,825
253	jamón cocido	0,00285	0,022	42,147	0,585
254	jamón ibérico	0,00531	0,044	42,191	1,170
255	jamón serrano	0,01324	0,088	42,278	2,339
256	jamón york	0,02178	0,131	42,410	3,509
257	JB	0,01038	0,044	42,453	1,170
258	jengibre	0,00214	0,022	42,475	0,585
259	jibia	0,00460	0,022	42,497	0,585
260	judía	0,04899	0,350	42,848	9,357
261	judía verde	0,00584	0,044	42,892	1,170
262	kebab	0,00460	0,022	42,913	0,585
263	ketchup	0,01751	0,175	43,089	4,678
264	kiko	0,00168	0,022	43,111	0,585
265	kiwi	0,01835	0,175	43,286	4,678
266	La Casera	0,00194	0,022	43,308	0,585
267	langostino	0,00532	0,044	43,352	1,170
268	lasaña	0,03895	0,241	43,593	6,433
269	láudano	0,00214	0,022	43,614	0,585
270	lechal	0,00214	0,022	43,636	0,585
271	leche	0,28277	1,555	45,192	41,520
272	leche condensada	0,00679	0,044	45,235	1,170
273	leche merengada	0,00613	0,044	45,279	1,170
274	lechuga	0,27775	1,818	47,097	48,538
275	legumbre	0,08494	0,526	47,623	14,035
276	lenguado	0,01619	0,088	47,711	2,339
277	lenteja	0,15435	0,942	48,653	25,146
278	licor	0,02035	0,153	48,806	4,094
279	lima	0,00482	0,044	48,850	1,170
280	limón	0,13157	0,942	49,792	25,146
281	limonada	0,03992	0,263	50,055	7,018
282	limoncello	0,00857	0,044	50,099	1,170
283	lomo	0,01808	0,131	50,230	3,509
284	lomo en manteca	0,00546	0,044	50,274	1,170
285	macarrón	0,09237	0,570	50,843	15,205
286	macedonia	0,00272	0,022	50,865	0,585

287	magdalena	0,01397	0,131	50,997	3,509
288	maíz	0,07650	0,526	51,522	14,035
289	mandarina	0,05121	0,394	51,917	10,526
290	mango	0,02580	0,197	52,114	5,263
291	manita de cerdo	0,00439	0,022	52,136	0,585
292	manteca	0,00761	0,044	52,180	1,170
293	mantecado	0,00185	0,022	52,202	0,585
294	mantel	0,00071	0,022	52,223	0,585
295	mantequilla	0,08239	0,723	52,946	19,298
296	manzana	0,18350	1,227	54,173	32,749
297	manzanilla	0,01564	0,088	54,261	2,339
298	maracuyá	0,00235	0,022	54,283	0,585
299	margarina	0,01594	0,153	54,436	4,094
300	marisco	0,02858	0,219	54,655	5,848
301	Martini	0,00362	0,022	54,677	0,585
302	mate	0,00235	0,022	54,699	0,585
303	mayonesa	0,01999	0,197	54,896	5,263
304	McDonald's	0,00460	0,022	54,918	0,585
305	mejillón	0,02312	0,175	55,093	4,678
306	melocotón	0,06460	0,504	55,597	13,450
307	melocotón en almíbar	0,00160	0,022	55,619	0,585
308	melón	0,10963	0,767	56,386	20,468
309	membrillo	0,00390	0,044	56,429	1,170
310	menta poleo	0,01338	0,088	56,517	2,339
311	merluza	0,00762	0,066	56,583	1,754
312	mermelada	0,04588	0,350	56,933	9,357
313	mero	0,00645	0,044	56,977	1,170
314	miel	0,01608	0,110	57,087	2,924
315	migas	0,00815	0,088	57,174	2,339
316	mijo	0,00460	0,022	57,196	0,585
317	mitad	0,00345	0,022	57,218	0,585
318	mojito	0,00418	0,022	57,240	0,585
319	mollete	0,00299	0,022	57,262	0,585
320	mora	0,01597	0,110	57,371	2,924
321	morcilla	0,02962	0,175	57,547	4,678
322	mortadela	0,00803	0,088	57,634	2,339
323	mortadela de aceitunas	0,00756	0,066	57,700	1,754
324	mostaza	0,00346	0,044	57,744	1,170
325	mozzarella	0,00329	0,022	57,766	0,585
326	muffin	0,00185	0,022	57,788	0,585
327	muslo	0,00299	0,022	57,809	0,585
328	nabo	0,00314	0,022	57,831	0,585
329	napolitana	0,00576	0,044	57,875	1,170
330	naranja	0,19793	1,336	59,211	35,673

331	naranjada	0,00285	0,022	59,233	0,585
332	nata	0,02476	0,197	59,430	5,263
333	natilla	0,01449	0,131	59,562	3,509
334	nectarina	0,01061	0,088	59,650	2,339
335	Nestea	0,01605	0,088	59,737	2,339
336	níspero	0,00912	0,088	59,825	2,339
337	nube	0,00920	0,044	59,869	1,170
338	nuez	0,01490	0,175	60,044	4,678
339	nugget	0,00329	0,022	60,066	0,585
340	ñoqui	0,00120	0,022	60,088	0,585
341	orégano	0,01290	0,110	60,197	2,924
342	ostra	0,00194	0,022	60,219	0,585
343	pacharán	0,00483	0,022	60,241	0,585
344	paella	0,11741	0,657	60,898	17,544
345	palmera	0,00483	0,022	60,920	0,585
346	palmito	0,00177	0,022	60,942	0,585
347	palomita	0,00299	0,022	60,964	0,585
348	pan	0,29139	1,709	62,673	45,614
349	pan cateto	0,00418	0,022	62,694	0,585
350	pan de ajo	0,00399	0,022	62,716	0,585
351	pan de molde	0,00533	0,044	62,760	1,170
352	pan integral	0,00513	0,044	62,804	1,170
353	panceta	0,00385	0,044	62,848	1,170
354	panini	0,01461	0,088	62,935	2,339
355	papaya	0,00473	0,044	62,979	1,170
356	papilla	0,00153	0,022	63,001	0,585
357	pasa	0,00214	0,022	63,023	0,585
358	pasta	0,17210	1,073	64,096	28,655
359	pastel	0,03758	0,307	64,403	8,187
360	patata	0,30074	1,884	66,287	50,292
361	patata a lo pobre	0,00224	0,022	66,309	0,585
362	patata asada	0,00829	0,066	66,375	1,754
363	patata frita	0,04131	0,241	66,616	6,433
364	paté	0,01506	0,131	66,747	3,509
365	pato	0,01414	0,088	66,835	2,339
366	pavo	0,03669	0,263	67,097	7,018
367	pechuga	0,00362	0,022	67,119	0,585
368	pepinillo	0,00817	0,088	67,207	2,339
369	pepino	0,05119	0,372	67,579	9,942
370	Pepsi	0,01142	0,044	67,623	1,170
371	pera	0,10017	0,701	68,324	18,713
372	perejil	0,02043	0,175	68,499	4,678
373	pero	0,00790	0,044	68,543	1,170
374	perrito (caliente)	0,00816	0,044	68,587	1,170
375	pescado	0,31706	1,862	70,449	49,708
376	pez espada	0,00968	0,088	70,537	2,339

377	picadillo	0,00194	0,022	70,559	0,585
378	picatoste	0,00818	0,066	70,624	1,754
379	pimentón	0,00331	0,044	70,668	1,170
380	pimienta	0,05350	0,416	71,084	11,111
381	pimiento	0,10175	0,767	71,851	20,468
382	pimiento morrón	0,00299	0,022	71,873	0,585
383	pinchito	0,00506	0,022	71,895	0,585
384	piña	0,06522	0,526	72,421	14,035
385	piñón	0,00126	0,022	72,442	0,585
386	pionono	0,00534	0,044	72,486	1,170
387	pipa	0,00345	0,044	72,530	1,170
388	pipa de calabaza	0,00177	0,022	72,552	0,585
389	pipa de girasol	0,00185	0,022	72,574	0,585
390	piquito	0,00185	0,022	72,596	0,585
391	pisto	0,00489	0,044	72,640	1,170
392	pitufio	0,00679	0,044	72,683	1,170
393	pizza	0,13286	0,854	73,538	22,807
394	plátano	0,12699	0,920	74,458	24,561
395	pollo	0,17838	1,139	75,597	30,409
396	pollo al curry	0,00329	0,022	75,619	0,585
397	pollo asado	0,01104	0,066	75,685	1,754
398	pomelo	0,00468	0,044	75,728	1,170
399	ponche	0,00362	0,022	75,750	0,585
400	porra	0,01097	0,066	75,816	1,754
401	porra antequerana	0,00795	0,066	75,882	1,754
402	postre	0,03404	0,263	76,145	7,018
403	potaje	0,09398	0,548	76,692	14,620
404	potaje de acelgas	0,00399	0,022	76,714	0,585
405	potaje de lentejas	0,00314	0,022	76,736	0,585
406	producto biológico	0,00235	0,022	76,758	0,585
407	puchero	0,07433	0,416	77,174	11,111
408	puerro	0,01036	0,088	77,262	2,339
409	Puleva	0,00259	0,022	77,284	0,585
410	pulpo	0,00532	0,044	77,327	1,170
411	puré	0,01359	0,153	77,481	4,094
412	puré de patata	0,00461	0,044	77,525	1,170
413	quesadilla	0,00168	0,022	77,547	0,585
414	queso	0,11047	0,898	78,445	23,977
415	queso azul	0,00377	0,044	78,488	1,170
416	queso blanco	0,00160	0,022	78,510	0,585
417	queso brie	0,00146	0,022	78,532	0,585
418	queso de untar	0,00168	0,022	78,554	0,585
419	queso manchego	0,00272	0,022	78,576	0,585
420	rábano	0,00557	0,044	78,620	1,170
421	rabo de toro	0,00177	0,022	78,642	0,585
422	rape	0,01107	0,066	78,708	1,754

423	rebanada	0,00177	0,022	78,729	0,585
424	rebujito	0,00418	0,022	78,751	0,585
425	receta	0,00185	0,022	78,773	0,585
426	Red Bull	0,00852	0,044	78,817	1,170
427	refresco	0,17408	0,942	79,759	25,146
428	remolacha	0,02769	0,219	79,978	5,848
429	repollo	0,00845	0,066	80,044	1,754
430	repostería	0,00168	0,022	80,066	0,585
431	restaurante	0,00194	0,022	80,088	0,585
432	revuelto	0,00214	0,022	80,110	0,585
433	revuelto de espárragos	0,00204	0,022	80,131	0,585
434	revuelto de verdura	0,00329	0,022	80,153	0,585
435	Ribera del Duero	0,00460	0,022	80,175	0,585
436	riñón al jerez	0,00247	0,022	80,197	0,585
437	rioja	0,00439	0,022	80,219	0,585
438	rollito de primavera	0,00247	0,022	80,241	0,585
439	rollo	0,00160	0,022	80,263	0,585
440	ron	0,05032	0,285	80,548	7,602
441	roquefort	0,00404	0,044	80,591	1,170
442	rosada	0,00927	0,066	80,657	1,754
443	ruca	0,00235	0,022	80,679	0,585
444	sacarina	0,00329	0,022	80,701	0,585
445	sake	0,00483	0,022	80,723	0,585
446	sal	0,09122	0,701	81,424	18,713
447	salami	0,00400	0,044	81,468	1,170
448	salchicha	0,04889	0,350	81,818	9,357
449	salchichón	0,02773	0,307	82,125	8,187
450	salmón	0,03050	0,197	82,322	5,263
451	salmonete	0,00146	0,022	82,344	0,585
452	salmorejo	0,01810	0,131	82,475	3,509
453	salsa	0,03567	0,263	82,738	7,018
454	salsa boloñesa	0,00362	0,022	82,760	0,585
455	salsa de yogur	0,00702	0,066	82,826	1,754
456	salsa rosa	0,00770	0,066	82,892	1,754
457	salteado	0,00126	0,022	82,913	0,585
458	sandía	0,08688	0,657	83,571	17,544
459	sándwich	0,03438	0,219	83,790	5,848
460	sangría	0,00966	0,066	83,855	1,754
461	sanjacobo	0,00547	0,066	83,921	1,754
462	sardina	0,01465	0,131	84,053	3,509
463	sauvignon	0,00531	0,022	84,074	0,585
464	secreto ibérico	0,00194	0,022	84,096	0,585
465	seso asado	0,00483	0,022	84,118	0,585
466	seta	0,01190	0,088	84,206	2,339
467	Seven Up	0,01635	0,088	84,294	2,339

468	Sex on the Beach	0,00380	0,022	84,315	0,585
469	shawarma	0,00439	0,022	84,337	0,585
470	sidra	0,00362	0,022	84,359	0,585
471	sirope	0,00185	0,022	84,381	0,585
472	sobrasada	0,00285	0,022	84,403	0,585
473	soda	0,00299	0,022	84,425	0,585
474	sofrito	0,00418	0,022	84,447	0,585
475	soja	0,01378	0,110	84,556	2,924
476	solomillo	0,04586	0,285	84,841	7,602
477	solomillo a la pimienta	0,00531	0,022	84,863	0,585
478	sopa	0,11987	0,701	85,564	18,713
479	sopa de marisco	0,00177	0,022	85,586	0,585
480	sopa de picadillo	0,00399	0,022	85,608	0,585
481	Sprite	0,00698	0,044	85,652	1,170
482	sushi	0,00967	0,044	85,696	1,170
483	tapa	0,00177	0,022	85,717	0,585
484	tarta	0,06275	0,438	86,156	11,696
485	tarta de chocolate	0,00506	0,022	86,177	0,585
486	tarta de fruta	0,00235	0,022	86,199	0,585
487	té	0,13456	0,789	86,988	21,053
488	té verde	0,00418	0,022	87,010	0,585
489	tenedor	0,00062	0,022	87,032	0,585
490	ternera	0,11876	0,811	87,842	21,637
491	tila	0,01536	0,088	87,930	2,339
492	tinto de verano	0,02391	0,131	88,061	3,509
493	tintorera	0,00439	0,022	88,083	0,585
494	tocino	0,00886	0,044	88,127	1,170
495	tocino de cielo	0,00185	0,022	88,149	0,585
496	tofu	0,00748	0,066	88,215	1,754
497	tomate	0,34845	2,278	90,493	60,819
498	tomate cherry	0,00224	0,022	90,515	0,585
499	tomate frito	0,01820	0,153	90,668	4,094
500	tomillo	0,00531	0,044	90,712	1,170
501	tónica	0,02122	0,153	90,865	4,094
502	tortellini	0,00418	0,022	90,887	0,585
503	tortilla	0,12264	0,767	91,654	20,468
504	tortilla de patatas	0,05125	0,263	91,917	7,018
505	tortilla francesa	0,01330	0,066	91,982	1,754
506	tostada	0,01052	0,088	92,070	2,339
507	trigo	0,00895	0,066	92,136	1,754
508	trucha	0,01106	0,066	92,202	1,754
509	uva	0,04029	0,350	92,552	9,357
510	vaca	0,00329	0,022	92,574	0,585
511	vainilla	0,00285	0,022	92,596	0,585
512	valeriana	0,00362	0,022	92,618	0,585

513	vegetal	0,01759	0,088	92,705	2,339
514	verdura	0,19704	1,227	93,932	32,749
515	vinagre	0,03181	0,285	94,217	7,602
516	vino	0,31504	1,752	95,969	46,784
517	vino blanco	0,01156	0,066	96,035	1,754
518	vino rosado	0,00659	0,044	96,079	1,170
519	vino tinto	0,02058	0,110	96,188	2,924
520	vodka	0,01831	0,131	96,320	3,509
521	wrap	0,00380	0,022	96,342	0,585
522	yogur	0,07681	0,657	96,999	17,544
523	zanahoria	0,13648	0,986	97,985	26,316
524	zummo	0,32295	1,774	99,759	47,368
525	zummo de arándano	0,00160	0,022	99,781	0,585
526	zummo de fruta	0,00553	0,044	99,825	1,170
527	zummo de melocotón	0,00177	0,022	99,847	0,585
528	zummo de naranja	0,00889	0,088	99,934	2,339
529	zummo de piña	0,00314	0,044	99,978	1,170
530	zurrapa	0,00194	0,022	100,000	0,585

1.1.4. La escuela (muebles y materiales)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(ordenador) portátil	0,05058	0,350	0,350	7,602
2	(rotulador) fluorescente	0,01023	0,108	0,457	2,339
3	abrigo	0,00224	0,027	0,484	0,585
4	actividad	0,00461	0,054	0,538	1,170
5	acuarela	0,00805	0,081	0,619	1,754
6	administrativo	0,00380	0,027	0,646	0,585
7	afilador	0,00204	0,027	0,672	0,585
8	agenda (escolar)	0,06635	0,511	1,183	11,111
9	agua	0,00194	0,027	1,210	0,585
10	aire acondicionado	0,00460	0,027	1,237	0,585
11	alfombra	0,00214	0,027	1,264	0,585
12	altavoz	0,01488	0,134	1,399	2,924
13	alumnado	0,00259	0,027	1,425	0,585
14	alumno	0,16535	1,076	2,501	23,392
15	amigo	0,00886	0,081	2,582	1,754
16	AMPA	0,00329	0,027	2,609	0,585
17	aprender	0,00204	0,027	2,636	0,585
18	aprendizaje	0,00285	0,027	2,663	0,585
19	aprobado	0,00345	0,027	2,690	0,585
20	aprobar	0,00235	0,027	2,717	0,585
21	apunte	0,01104	0,108	2,824	2,339
22	archivador	0,02605	0,269	3,093	5,848
23	armario	0,14676	1,049	4,142	22,807
24	aro	0,00139	0,027	4,169	0,585
25	ascensor	0,00329	0,027	4,196	0,585
26	aseo	0,01851	0,134	4,330	2,924
27	asignatura	0,01780	0,134	4,465	2,924
28	atlas	0,00546	0,054	4,519	1,170
29	atril	0,01329	0,108	4,626	2,339
30	aula	0,07924	0,565	5,191	12,281
31	aula de convivencia	0,00362	0,027	5,218	0,585
32	auricular	0,00185	0,027	5,245	0,585
33	autobús (escolar)	0,00585	0,054	5,299	1,170
34	babero	0,00139	0,027	5,325	0,585
35	baldosa	0,00329	0,027	5,352	0,585
36	balón	0,01803	0,188	5,541	4,094
37	banca	0,00664	0,054	5,594	1,170
38	banco	0,04444	0,323	5,917	7,018
39	banqueta	0,00878	0,054	5,971	1,170
40	barro	0,00259	0,027	5,998	0,585
41	Biblia	0,00285	0,027	6,025	0,585
42	biblioteca	0,07941	0,699	6,724	15,205
43	bigotera	0,00299	0,027	6,751	0,585

44	biología	0,00561	0,054	6,805	1,170
45	bloc	0,00299	0,027	6,832	0,585
46	bloc de dibujo	0,00168	0,027	6,859	0,585
47	bloc de notas	0,00531	0,027	6,885	0,585
48	boca de incendio equipada	0,00380	0,027	6,912	0,585
49	bocadillo	0,00820	0,054	6,966	1,170
50	bola del mundo	0,00299	0,027	6,993	0,585
51	boli(grafo)	0,58391	3,846	10,839	83,626
52	bolso	0,00285	0,027	10,866	0,585
53	bombilla	0,00506	0,027	10,893	0,585
54	borrador	0,26026	1,694	12,587	36,842
55	butaca	0,00177	0,027	12,614	0,585
56	cable	0,00177	0,027	12,641	0,585
57	cafetería	0,00503	0,054	12,695	1,170
58	caja	0,00490	0,054	12,749	1,170
59	cajón	0,01107	0,108	12,856	2,339
60	cajonera	0,00483	0,027	12,883	0,585
61	calculadora	0,08936	0,807	13,690	17,544
62	calendario	0,01330	0,134	13,825	2,924
63	caligrafía	0,00194	0,027	13,852	0,585
64	cama	0,00380	0,027	13,878	0,585
65	campana	0,00439	0,027	13,905	0,585
66	campus	0,00214	0,027	13,932	0,585
67	canasta	0,00478	0,054	13,986	1,170
68	cancha de baloncesto	0,00966	0,081	14,067	1,754
69	cantina	0,00259	0,027	14,094	0,585
70	cañón	0,05671	0,430	14,524	9,357
71	capilla	0,00259	0,027	14,551	0,585
72	carpeta	0,16035	1,479	16,030	32,164
73	cartabón	0,10464	0,861	16,891	18,713
74	cartera	0,01622	0,161	17,052	3,509
75	cartilla	0,00204	0,027	17,079	0,585
76	cartuchera	0,01767	0,134	17,214	2,924
77	cartulina	0,05400	0,565	17,778	12,281
78	castigo	0,00764	0,081	17,859	1,754
79	CD(-ROM)	0,00863	0,081	17,940	1,754
80	celo	0,00509	0,054	17,994	1,170
81	celofán	0,00160	0,027	18,020	0,585
82	centro bilingüe	0,00329	0,027	18,047	0,585
83	centro TIC	0,00904	0,081	18,128	1,754
84	chándal	0,00329	0,027	18,155	0,585
85	chincheta	0,01064	0,134	18,289	2,924
86	ciencias naturales	0,00247	0,027	18,316	0,585
87	clase	0,11999	0,807	19,123	17,544
88	clasificador	0,03539	0,296	19,419	6,433
89	claustro	0,01428	0,081	19,500	1,754
90	clip	0,01651	0,188	19,688	4,094

91	colchoneta	0,00758	0,081	19,769	1,754
92	colegio	0,01522	0,108	19,876	2,339
93	color	0,03888	0,377	20,253	8,187
94	columpio	0,00235	0,027	20,280	0,585
95	comedor	0,01889	0,188	20,468	4,094
96	compañero	0,01576	0,108	20,576	2,339
97	compás	0,14419	1,291	21,867	28,070
98	comportamiento	0,00214	0,027	21,893	0,585
99	concurso	0,00314	0,027	21,920	0,585
100	conocimiento	0,00160	0,027	21,947	0,585
101	consejo escolar	0,00345	0,027	21,974	0,585
102	conserje	0,03366	0,269	22,243	5,848
103	conserjería	0,03644	0,323	22,566	7,018
104	control	0,00451	0,054	22,620	1,170
105	copistería	0,00937	0,081	22,700	1,754
106	corcho	0,02221	0,215	22,916	4,678
107	corrector	0,01239	0,108	23,023	2,339
108	cortina	0,00380	0,027	23,050	0,585
109	cuadernillo	0,01318	0,134	23,185	2,924
110	cuaderno	0,23802	1,802	24,987	39,181
111	cuadrícula	0,00214	0,027	25,013	0,585
112	cuadro	0,00299	0,027	25,040	0,585
113	cuerda	0,00160	0,027	25,067	0,585
114	cura	0,00224	0,027	25,094	0,585
115	curso académico	0,00247	0,027	25,121	0,585
116	deber	0,02061	0,215	25,336	4,678
117	departamento	0,00910	0,081	25,417	1,754
118	desayuno	0,00120	0,027	25,444	0,585
119	despacho	0,00292	0,054	25,498	1,170
120	diapositiva	0,00571	0,054	25,551	1,170
121	diario	0,00153	0,027	25,578	0,585
122	dibujo	0,00833	0,081	25,659	1,754
123	diccionario	0,06753	0,646	26,304	14,035
124	dintel	0,00439	0,027	26,331	0,585
125	dirección	0,00329	0,027	26,358	0,585
126	director	0,02985	0,269	26,627	5,848
127	diversión	0,00204	0,027	26,654	0,585
128	dossier	0,00259	0,027	26,681	0,585
129	edificio	0,00362	0,027	26,708	0,585
130	educación física	0,00224	0,027	26,735	0,585
131	educación plástica y visual	0,00449	0,054	26,789	1,170
132	ejercicio	0,00549	0,054	26,842	1,170
133	encerado	0,01498	0,108	26,950	2,339
134	enchufe	0,00781	0,081	27,031	1,754
135	enciclopedia	0,01338	0,108	27,138	2,339
136	enemigo	0,00247	0,027	27,165	0,585
137	enseñar	0,00214	0,027	27,192	0,585
138	escalera	0,00457	0,054	27,246	1,170

139	escalímetro	0,00272	0,027	27,273	0,585
140	escáner	0,00204	0,027	27,300	0,585
141	escritorio	0,08527	0,457	27,757	9,942
142	escuadra	0,10981	0,888	28,644	19,298
143	escuela	0,01429	0,081	28,725	1,754
144	escuela concertada	0,00299	0,027	28,752	0,585
145	escuela privada	0,00285	0,027	28,779	0,585
146	escuela pública	0,00314	0,027	28,806	0,585
147	espejo	0,00259	0,027	28,833	0,585
148	esquema	0,00362	0,027	28,860	0,585
149	estante	0,00822	0,054	28,913	1,170
150	estantería	0,08346	0,592	29,505	12,865
151	estrado	0,01057	0,081	29,586	1,754
152	estuche	0,27066	2,205	31,791	47,953
153	estudiante	0,01986	0,134	31,926	2,924
154	estudiar	0,01189	0,108	32,033	2,339
155	estudio	0,00224	0,027	32,060	0,585
156	ética	0,00153	0,027	32,087	0,585
157	examen	0,03711	0,350	32,437	7,602
158	excursión	0,00259	0,027	32,464	0,585
159	extintor	0,00658	0,054	32,517	1,170
160	fax	0,00345	0,027	32,544	0,585
161	filosofía	0,00259	0,027	32,571	0,585
162	física	0,00799	0,108	32,679	2,339
163	Fixo	0,00439	0,027	32,706	0,585
164	folio	0,16718	1,452	34,158	31,579
165	forro	0,00399	0,027	34,185	0,585
166	fotocopia	0,00880	0,108	34,293	2,339
167	fotocopiadora	0,01158	0,134	34,427	2,924
168	francés	0,00345	0,027	34,454	0,585
169	fuelle	0,00185	0,027	34,481	0,585
170	funda	0,00460	0,027	34,508	0,585
171	fútbol	0,00314	0,027	34,535	0,585
172	gafa	0,00607	0,081	34,615	1,754
173	geografía	0,00259	0,027	34,642	0,585
174	geología	0,00299	0,027	34,669	0,585
175	gimnasia	0,00617	0,081	34,750	1,754
176	gimnasio	0,06358	0,511	35,261	11,111
177	globo terráqueo	0,00673	0,081	35,342	1,754
178	goma (de borrar)	0,41416	2,932	38,273	63,743
179	grande	0,00194	0,027	38,300	0,585
180	grapa	0,00732	0,054	38,354	1,170
181	grapadora	0,03779	0,350	38,704	7,602
182	griego	0,00272	0,027	38,731	0,585
183	grupo	0,00132	0,027	38,757	0,585
184	guardería	0,00214	0,027	38,784	0,585
185	guitarra	0,00109	0,027	38,811	0,585
186	hall	0,00247	0,027	38,838	0,585
187	historia	0,00432	0,054	38,892	1,170

188	hoja	0,03300	0,269	39,161	5,848
189	horario	0,02901	0,269	39,430	5,848
190	iluminación	0,00439	0,027	39,457	0,585
191	imagen	0,00204	0,027	39,484	0,585
192	imbornal	0,00460	0,027	39,510	0,585
193	impresora	0,00235	0,027	39,537	0,585
194	inglés	0,00362	0,027	39,564	0,585
195	institución	0,00380	0,027	39,591	0,585
196	instituto	0,00767	0,054	39,645	1,170
197	internet	0,02515	0,242	39,887	5,263
198	jardín	0,00272	0,027	39,914	0,585
199	jefatura de estudios	0,00460	0,027	39,941	0,585
200	jefe de estudios	0,01833	0,134	40,075	2,924
201	juego	0,00868	0,108	40,183	2,339
202	juguete	0,00961	0,108	40,290	2,339
203	laboratorio	0,01140	0,108	40,398	2,339
204	lámina	0,00627	0,054	40,452	1,170
205	lámpara	0,04037	0,296	40,748	6,433
206	lapicero	0,03420	0,269	41,017	5,848
207	lápiz	0,62040	4,008	45,024	87,135
208	lápiz de cera	0,06255	0,538	45,562	11,696
209	lápiz de color	0,06411	0,511	46,073	11,111
210	lápiz de madera	0,00460	0,027	46,100	0,585
211	latín	0,00285	0,027	46,127	0,585
212	lavabo	0,00571	0,054	46,181	1,170
213	lectura	0,00204	0,027	46,208	0,585
214	lengua	0,00621	0,081	46,288	1,754
215	lengua castellana y literatura	0,00329	0,027	46,315	0,585
216	lenguaje	0,00259	0,027	46,342	0,585
217	librería	0,00224	0,027	46,369	0,585
218	libreta	0,29204	2,152	48,521	46,784
219	libro	0,46042	3,416	51,937	74,269
220	libro de lectura	0,00329	0,027	51,963	0,585
221	libro de texto	0,01777	0,108	52,071	2,339
222	libro del profesor	0,00177	0,027	52,098	0,585
223	líquido corrector	0,00214	0,027	52,125	0,585
224	lista	0,00567	0,054	52,179	1,170
225	luz	0,02017	0,188	52,367	4,094
226	madrugar	0,00120	0,027	52,394	0,585
227	maestro	0,01242	0,081	52,474	1,754
228	maleta	0,04922	0,377	52,851	8,187
229	manual	0,00285	0,027	52,878	0,585
230	manualidad	0,00259	0,027	52,905	0,585
231	mapa	0,05169	0,484	53,389	10,526
232	mapamundi	0,00345	0,027	53,416	0,585
233	marcador	0,00362	0,027	53,443	0,585
234	matemáticas	0,01232	0,134	53,577	2,924
235	materia	0,00418	0,027	53,604	0,585

236	material escolar	0,01323	0,081	53,685	1,754
237	matrícula	0,00224	0,027	53,712	0,585
238	mesa	0,66115	3,470	57,181	75,439
239	mesa de dibujo	0,00235	0,027	57,208	0,585
240	mesa del profesor	0,01197	0,081	57,289	1,754
241	método	0,00272	0,027	57,316	0,585
242	mina	0,00945	0,081	57,396	1,754
243	mochila	0,22055	1,910	59,306	41,520
244	monja	0,00362	0,027	59,333	0,585
245	muñeco	0,00146	0,027	59,360	0,585
246	mural	0,00399	0,027	59,387	0,585
247	música	0,00343	0,054	59,441	1,170
248	netbook	0,00362	0,027	59,467	0,585
249	niño	0,00460	0,054	59,521	1,170
250	nota	0,00689	0,081	59,602	1,754
251	ordenador	0,35321	2,582	62,184	56,140
252	ordenador personal	0,00399	0,027	62,211	0,585
253	orientador	0,00224	0,027	62,238	0,585
254	página	0,00418	0,027	62,265	0,585
255	pandereta	0,00115	0,027	62,292	0,585
256	panel	0,00314	0,027	62,318	0,585
257	pantalla	0,05684	0,430	62,749	9,357
258	pantalla digital	0,00585	0,027	62,776	0,585
259	papel	0,11556	0,941	63,717	20,468
260	papel cebolla	0,00139	0,027	63,744	0,585
261	papel de celofán	0,01069	0,081	63,825	1,754
262	papel de seda	0,00247	0,027	63,852	0,585
263	papel milimetrado	0,00153	0,027	63,878	0,585
264	papel pinocho	0,00483	0,054	63,932	1,170
265	papel vegetal	0,00146	0,027	63,959	0,585
266	papelera	0,07313	0,646	64,605	14,035
267	paragüero	0,00272	0,027	64,632	0,585
268	pared	0,01527	0,108	64,739	2,339
269	parte	0,00445	0,054	64,793	1,170
270	parte de faltas	0,00185	0,027	64,820	0,585
271	pasillo	0,03526	0,242	65,062	5,263
272	patio (de recreo)	0,11472	0,914	65,976	19,883
273	pegamento	0,05809	0,511	66,487	11,111
274	pegatina	0,00235	0,027	66,514	0,585
275	película	0,00153	0,027	66,541	0,585
276	pelota	0,00658	0,054	66,595	1,170
277	pendrive	0,00299	0,027	66,622	0,585
278	percha	0,01759	0,134	66,756	2,924
279	perchero	0,07641	0,619	67,375	13,450
280	persiana	0,01267	0,134	67,509	2,924
281	personal de limpieza	0,00235	0,027	67,536	0,585

282	pica	0,00153	0,027	67,563	0,585
283	Pilot	0,00483	0,027	67,590	0,585
284	pincel	0,00981	0,108	67,698	2,339
285	pintura	0,01363	0,134	67,832	2,924
286	pista deportiva	0,01133	0,108	67,940	2,339
287	pizarra	0,75097	4,061	72,001	88,304
288	pizarra blanca	0,00531	0,027	72,028	0,585
289	pizarra digital	0,05500	0,323	72,351	7,018
290	pizarra tradicional	0,00531	0,027	72,378	0,585
291	plastilina	0,02555	0,269	72,647	5,848
292	pluma	0,05029	0,350	72,996	7,602
293	plumier	0,00247	0,027	73,023	0,585
294	polideportivo	0,00247	0,027	73,050	0,585
295	porta-ángulos	0,00380	0,027	73,077	0,585
296	portafolio	0,00204	0,027	73,104	0,585
297	portaminas	0,05487	0,430	73,534	9,357
298	portería	0,00146	0,027	73,561	0,585
299	portero	0,00214	0,027	73,588	0,585
300	póster	0,01987	0,188	73,776	4,094
301	Post-it	0,01798	0,161	73,938	3,509
302	potro	0,00247	0,027	73,964	0,585
303	probeta	0,00204	0,027	73,991	0,585
304	producto químico	0,00247	0,027	74,018	0,585
305	profesor	0,22485	1,399	75,417	30,409
306	profesorado	0,00272	0,027	75,444	0,585
307	programación	0,00247	0,027	75,471	0,585
308	proyector	0,19520	1,399	76,869	30,409
309	puerta	0,15163	1,076	77,945	23,392
310	pupilo	0,00531	0,027	77,972	0,585
311	pupitre	0,41123	2,205	80,178	47,953
312	química	0,00538	0,081	80,258	1,754
313	quitagrapas	0,00329	0,027	80,285	0,585
314	radio	0,01961	0,188	80,473	4,094
315	rampa para minusválidos	0,00314	0,027	80,500	0,585
316	raqueta	0,00153	0,027	80,527	0,585
317	ratón	0,01605	0,108	80,635	2,339
318	recreo	0,06301	0,538	81,173	11,696
319	redacción	0,00153	0,027	81,200	0,585
320	regla	0,27437	2,071	83,271	45,029
321	religión	0,00235	0,027	83,297	0,585
322	reloj	0,00272	0,027	83,324	0,585
323	repaso	0,00329	0,027	83,351	0,585
324	reproductor	0,00532	0,054	83,405	1,170
325	reproductor de DVD	0,00132	0,027	83,432	0,585

326	reproductor de vídeo	0,00399	0,027	83,459	0,585
327	respaldo	0,00483	0,027	83,486	0,585
328	resumen	0,00160	0,027	83,513	0,585
329	retroproyector	0,02210	0,215	83,728	4,678
330	reunión	0,00204	0,027	83,755	0,585
331	revista	0,00272	0,027	83,782	0,585
332	ropero	0,00380	0,027	83,809	0,585
333	rotulador	0,24592	1,910	85,718	41,520
334	sacapuntas	0,15239	1,130	86,848	24,561
335	sala de informática	0,00770	0,081	86,928	1,754
336	sala de música	0,00235	0,027	86,955	0,585
337	sala de ordenadores	0,01242	0,108	87,063	2,339
338	sala de profesores	0,02213	0,188	87,251	4,094
339	salón de actos	0,02429	0,215	87,466	4,678
340	sanción	0,00177	0,027	87,493	0,585
341	secretaría	0,01769	0,161	87,655	3,509
342	secretario	0,00633	0,054	87,708	1,170
343	semicírculo				
	graduado	0,00272	0,027	87,735	0,585
344	seminario	0,00506	0,027	87,762	0,585
345	Séneca	0,00194	0,027	87,789	0,585
346	separador	0,00177	0,027	87,816	0,585
347	servicio	0,01319	0,108	87,924	2,339
348	silla	0,68505	3,792	91,716	82,456
349	sillón	0,00830	0,081	91,797	1,754
350	subrayador	0,02637	0,215	92,012	4,678
351	suelo	0,01073	0,081	92,093	1,754
352	suspender	0,00247	0,027	92,119	0,585
353	suspenso	0,00362	0,027	92,146	0,585
354	tablón	0,03631	0,296	92,442	6,433
355	tablón de anuncios	0,01339	0,108	92,550	2,339
356	tablón de corcho	0,00272	0,027	92,577	0,585
357	taburete	0,00728	0,054	92,630	1,170
358	taller	0,00703	0,054	92,684	1,170
359	tambor	0,00126	0,027	92,711	0,585
360	taquilla	0,02796	0,269	92,980	5,848
361	tarea	0,00194	0,027	93,007	0,585
362	tarima	0,04675	0,323	93,330	7,018
363	techo	0,01423	0,108	93,437	2,339
364	teclado	0,02253	0,161	93,599	3,509
365	tecnología	0,00247	0,027	93,626	0,585
366	teléfono	0,00345	0,027	93,653	0,585
367	televisión	0,03041	0,269	93,921	5,848
368	televisor	0,00380	0,027	93,948	0,585
369	temario	0,00345	0,027	93,975	0,585

370	témpera	0,00942	0,108	94,083	2,339
371	tenis	0,00314	0,027	94,110	0,585
372	test	0,00299	0,027	94,137	0,585
373	tijera	0,06003	0,565	94,701	12,281
374	timbre	0,01284	0,134	94,836	2,924
375	tinta	0,00834	0,108	94,944	2,339
376	Tipp-Ex	0,06383	0,511	95,455	11,111
377	tiza	0,45688	2,797	98,252	60,819
378	trabajo	0,00204	0,027	98,279	0,585
379	transparencia	0,00544	0,054	98,332	1,170
380	transportador	0,00380	0,027	98,359	0,585
381	transportador de ángulos	0,00299	0,027	98,386	0,585
382	transporte escolar	0,00299	0,027	98,413	0,585
383	trimestre	0,00418	0,027	98,440	0,585
384	tubo fluorescente	0,00224	0,027	98,467	0,585
385	tutor	0,00185	0,027	98,494	0,585
386	unidad didáctica	0,00259	0,027	98,521	0,585
387	uniforme	0,01287	0,134	98,655	2,924
388	universidad	0,00731	0,054	98,709	1,170
389	urinario	0,00506	0,027	98,736	0,585
390	váter	0,00285	0,027	98,763	0,585
391	ventana	0,15467	1,183	99,946	25,731
392	vídeo	0,00779	0,054	100,000	1,170

1.1.5. La ciudad

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(auto)bus	0,23166	1,645	1,645	35,673
2	abuelo	0,00272	0,027	1,672	0,585
3	academia	0,01229	0,108	1,780	2,339
4	accidente	0,00790	0,081	1,861	1,754
5	acequia	0,00506	0,027	1,888	0,585
6	acera	0,19128	1,375	3,263	29,825
7	acueducto	0,00299	0,027	3,290	0,585
8	adosado	0,01146	0,108	3,398	2,339
9	aduanas	0,00483	0,027	3,425	0,585
10	aeropuerto	0,11264	0,890	4,315	19,298
11	afueras	0,00214	0,027	4,342	0,585
12	agencia	0,00399	0,027	4,369	0,585
13	agencia de viajes	0,00709	0,054	4,423	1,170
14	aglomeración	0,01381	0,108	4,531	2,339
15	ajeteo	0,00950	0,081	4,612	1,754
16	alameda	0,00460	0,027	4,639	0,585
17	albergue	0,00519	0,054	4,693	1,170
18	alcalde	0,01116	0,054	4,746	1,170
19	alcantarilla	0,00996	0,081	4,827	1,754
20	alcantarillado	0,00285	0,027	4,854	0,585
21	alcazaba	0,01902	0,135	4,989	2,924
22	Alliance Française	0,00483	0,027	5,016	0,585
23	almacén	0,01164	0,081	5,097	1,754
24	alumbrado	0,00314	0,027	5,124	0,585
25	ambulancia	0,00933	0,081	5,205	1,754
26	ambulatorio	0,00194	0,027	5,232	0,585
27	amplitud	0,00120	0,027	5,259	0,585
28	anciano	0,00576	0,054	5,313	1,170
29	andén	0,00314	0,027	5,340	0,585
30	animal	0,00177	0,027	5,367	0,585
31	ansiedad	0,00362	0,027	5,394	0,585
32	antena	0,00329	0,054	5,448	1,170
33	anuncio	0,00670	0,054	5,502	1,170
34	aparcamiento	0,05957	0,512	6,014	11,111
35	apartamento	0,03186	0,243	6,257	5,263
36	árbol	0,14364	1,052	7,309	22,807
37	arbusto	0,00214	0,027	7,335	0,585
38	arcén	0,00830	0,054	7,389	1,170
39	arriate	0,00204	0,027	7,416	0,585
40	ascensor	0,01207	0,081	7,497	1,754
41	asfalto	0,01464	0,108	7,605	2,339
42	asiento	0,00362	0,027	7,632	0,585
43	atasco	0,04223	0,297	7,929	6,433
44	ático	0,00659	0,054	7,983	1,170

45	autobús urbano	0,00247	0,027	8,010	0,585
46	autopista	0,02141	0,135	8,145	2,924
47	autovía	0,03180	0,216	8,360	4,678
48	AVE	0,00652	0,054	8,414	1,170
49	avenida	0,18344	1,133	9,547	24,561
50	avión	0,01827	0,108	9,655	2,339
51	ayuntamiento	0,18215	1,268	10,922	27,485
52	badén	0,00285	0,027	10,949	0,585
53	balcón	0,00132	0,027	10,976	0,585
54	baldosa	0,00407	0,054	11,030	1,170
55	banco	0,15364	1,241	12,271	26,901
56	bandera	0,00314	0,027	12,298	0,585
57	bar	0,18353	1,456	13,754	31,579
58	bar de tapas	0,00247	0,027	13,781	0,585
59	barbería	0,00259	0,027	13,808	0,585
60	Barcelona	0,00314	0,027	13,835	0,585
61	barco	0,00460	0,027	13,862	0,585
62	barrendero	0,00204	0,027	13,889	0,585
63	barrio	0,07248	0,566	14,455	12,281
64	basílica	0,00506	0,027	14,482	0,585
65	basura	0,02252	0,189	14,671	4,094
66	bazar	0,00380	0,027	14,698	0,585
67	biblioteca	0,04662	0,405	15,102	8,772
68	bici(cleta)	0,07485	0,512	15,615	11,111
69	bloque	0,02173	0,135	15,750	2,924
70	bloque de pisos	0,02715	0,189	15,939	4,094
71	boca de incendio	0,00345	0,027	15,965	0,585
72	boca de metro	0,00483	0,027	15,992	0,585
73	bolera	0,00329	0,027	16,019	0,585
74	Bolsa	0,00224	0,027	16,046	0,585
75	bombero	0,01215	0,108	16,154	2,339
76	bordillo	0,01015	0,081	16,235	1,754
77	boutique	0,00168	0,027	16,262	0,585
78	bufete de abogados	0,01221	0,081	16,343	1,754
79	bulevar	0,03411	0,216	16,559	4,678
80	bullicio	0,01083	0,081	16,640	1,754
81	Burger King	0,00362	0,027	16,667	0,585
82	buzón	0,00483	0,027	16,694	0,585
83	cabina	0,01100	0,081	16,775	1,754
84	cabina telefónica	0,00878	0,054	16,828	1,170
85	cable	0,00153	0,027	16,855	0,585
86	cafetería	0,07024	0,539	17,395	11,696
87	cajero (automático)	0,02041	0,162	17,557	3,509
88	cajero (automático)				
	señal	0,00418	0,027	17,584	0,585
89	calle	0,54776	2,940	20,523	63,743
90	calle peatonal	0,00711	0,081	20,604	1,754
91	callejón	0,03250	0,216	20,820	4,678
92	callejón sin salida	0,00483	0,027	20,847	0,585

93	callejuela	0,00886	0,054	20,901	1,170
94	calzada	0,01370	0,135	21,036	2,924
95	camino	0,00285	0,027	21,063	0,585
96	camión	0,02834	0,189	21,251	4,094
97	camión de la basura	0,00345	0,027	21,278	0,585
98	campo	0,00329	0,027	21,305	0,585
99	campo de fútbol	0,01531	0,135	21,440	2,924
100	campus (universitario)	0,00778	0,054	21,494	1,170
101	cansancio	0,00139	0,027	21,521	0,585
102	capilla	0,00859	0,054	21,575	1,170
103	capital	0,00272	0,027	21,602	0,585
104	caravana	0,01328	0,108	21,710	2,339
105	cárcel	0,00955	0,108	21,818	2,339
106	carnicería	0,04673	0,405	22,222	8,772
107	carretera	0,29351	1,807	24,029	39,181
108	carril	0,01092	0,081	24,110	1,754
109	carril bici	0,02609	0,189	24,299	4,094
110	carril bus	0,00630	0,081	24,380	1,754
111	cartel	0,00803	0,054	24,434	1,170
112	cartel publicitario	0,00576	0,054	24,488	1,170
113	casa	0,27729	1,807	26,294	39,181
114	casa adosada	0,00531	0,054	26,348	1,170
115	casa hermandad	0,00299	0,027	26,375	0,585
116	casa mata	0,00837	0,081	26,456	1,754
117	casco antiguo	0,01576	0,081	26,537	1,754
118	casco histórico	0,01794	0,108	26,645	2,339
119	castillo	0,00903	0,081	26,726	1,754
120	catedral	0,08922	0,647	27,373	14,035
121	ceda el paso	0,00380	0,027	27,400	0,585
122	cementerio	0,01251	0,108	27,508	2,339
123	centro	0,05882	0,378	27,886	8,187
124	centro cívico	0,00531	0,027	27,913	0,585
125	centro comercial	0,23725	1,780	29,693	38,596
126	centro concertado	0,00345	0,027	29,720	0,585
127	centro de belleza	0,00259	0,027	29,746	0,585
128	centro de estética	0,00380	0,027	29,773	0,585
129	centro de salud	0,01831	0,189	29,962	4,094
130	centro de transporte	0,01013	0,054	30,016	1,170
131	centro diurno	0,00314	0,027	30,043	0,585
132	centro histórico	0,02132	0,135	30,178	2,924
133	centro social	0,00329	0,027	30,205	0,585
134	Cercanías	0,00204	0,027	30,232	0,585
135	cervecería	0,00235	0,027	30,259	0,585
136	césped	0,00214	0,027	30,286	0,585
137	chabola	0,00153	0,027	30,313	0,585
138	chalé	0,01482	0,162	30,475	3,509
139	charcutería	0,00956	0,054	30,529	1,170
140	chimenea	0,00177	0,027	30,556	0,585

141	chiringuito	0,01381	0,135	30,690	2,924
142	cibercafé	0,00399	0,027	30,717	0,585
143	ciclista	0,00439	0,027	30,744	0,585
144	cielo	0,00399	0,027	30,771	0,585
145	cine	0,17308	1,321	32,093	28,655
146	circulación	0,00642	0,054	32,147	1,170
147	circunvalación	0,00418	0,027	32,174	0,585
148	ciudad	0,00329	0,027	32,201	0,585
149	ciudadano	0,00792	0,054	32,255	1,170
150	claxon	0,00460	0,027	32,282	0,585
151	clínica	0,00812	0,054	32,335	1,170
152	clínica privada	0,00362	0,027	32,362	0,585
153	clínica veterinaria	0,00439	0,027	32,389	0,585
154	club	0,00224	0,027	32,416	0,585
155	club nocturno	0,00418	0,027	32,443	0,585
156	coche	0,50193	2,913	35,356	63,158
157	coche de caballos	0,00362	0,027	35,383	0,585
158	colegio	0,13254	1,025	36,408	22,222
159	columpio	0,02166	0,189	36,597	4,094
160	comercio	0,05832	0,458	37,055	9,942
161	comisaría (de policía)	0,05802	0,485	37,540	10,526
162	compra	0,00601	0,054	37,594	1,170
163	comunidad	0,00259	0,027	37,621	0,585
164	concejal	0,00272	0,027	37,648	0,585
165	concierto	0,00380	0,027	37,675	0,585
166	conductor	0,00299	0,027	37,702	0,585
167	confitería	0,01168	0,108	37,810	2,339
168	conservatorio	0,00439	0,027	37,837	0,585
169	construcción	0,00531	0,027	37,864	0,585
170	consulado	0,00584	0,054	37,918	1,170
171	consultorio médico	0,00798	0,054	37,972	1,170
172	contador de agua	0,00235	0,027	37,999	0,585
173	contaminación	0,07056	0,512	38,511	11,111
174	contenedor	0,00362	0,027	38,538	0,585
175	contenedor de basura	0,01452	0,108	38,646	2,339
176	contenedor de reciclaje	0,00908	0,081	38,727	1,754
177	copistería	0,01177	0,108	38,835	2,339
178	cornisa	0,00380	0,027	38,862	0,585
179	Correos	0,03585	0,243	39,105	5,263
180	cosa	0,00204	0,027	39,132	0,585
181	cruce	0,02153	0,162	39,293	3,509
182	cuadro	0,00483	0,027	39,320	0,585
183	cuartel	0,00126	0,027	39,347	0,585
184	cuartel de la Guardia Civil	0,00329	0,027	39,374	0,585
185	cultivo	0,00247	0,027	39,401	0,585

186	cultura	0,00235	0,027	39,428	0,585
187	decomisos	0,00314	0,027	39,455	0,585
188	Delegación de Hacienda	0,00887	0,081	39,536	1,754
189	delegación ministerial	0,00299	0,027	39,563	0,585
190	desagüe	0,00299	0,027	39,590	0,585
191	dinero	0,00146	0,027	39,617	0,585
192	dirección	0,00314	0,027	39,644	0,585
193	discoteca	0,06725	0,593	40,237	12,865
194	diversión	0,00877	0,054	40,291	1,170
195	droguería	0,00329	0,027	40,318	0,585
196	dúplex	0,00506	0,027	40,345	0,585
197	edificio	0,49611	3,020	43,366	65,497
198	El Retiro	0,00224	0,027	43,393	0,585
199	electricidad	0,00259	0,027	43,420	0,585
200	Emasa	0,00194	0,027	43,447	0,585
201	embajada	0,00584	0,054	43,501	1,170
202	empresa	0,00285	0,027	43,528	0,585
203	EMT	0,00418	0,027	43,554	0,585
204	Endesa	0,00185	0,027	43,581	0,585
205	energía	0,00418	0,027	43,608	0,585
206	ente financiera	0,00247	0,027	43,635	0,585
207	entretenimiento	0,00362	0,027	43,662	0,585
208	escalera	0,00345	0,027	43,689	0,585
209	escalera automática	0,01003	0,081	43,770	1,754
210	escalinata	0,00483	0,027	43,797	0,585
211	escaparate	0,01766	0,135	43,932	2,924
212	escuela	0,09205	0,728	44,660	15,789
213	Escuela Oficial de Idiomas	0,01017	0,081	44,741	1,754
214	esquina	0,00531	0,027	44,768	0,585
215	establecimiento	0,00380	0,027	44,795	0,585
216	estación	0,07565	0,593	45,388	12,865
217	estación de autobús	0,02877	0,216	45,604	4,678
218	estación de metro	0,00698	0,054	45,658	1,170
219	estación de tren	0,07837	0,566	46,224	12,281
220	estadio	0,01718	0,189	46,413	4,094
221	estadio de fútbol	0,00380	0,027	46,440	0,585
222	estanco	0,00609	0,054	46,494	1,170
223	estanque	0,00299	0,027	46,521	0,585
224	estatua	0,03672	0,270	46,791	5,848
225	estrés	0,03086	0,189	46,980	4,094
226	estudio	0,00272	0,027	47,006	0,585
227	estudio fotográfico	0,00299	0,027	47,033	0,585
228	excavadora	0,00399	0,027	47,060	0,585
229	fábrica	0,00569	0,081	47,141	1,754
230	facultad	0,01364	0,135	47,276	2,924
231	farmacia	0,03832	0,297	47,573	6,433

232	faro	0,00299	0,027	47,600	0,585
233	farola	0,10210	0,728	48,328	15,789
234	ferretería	0,01403	0,135	48,463	2,924
235	festival	0,00259	0,027	48,490	0,585
236	fiesta	0,01114	0,081	48,571	1,754
237	flor	0,00768	0,054	48,625	1,170
238	floristería	0,01490	0,135	48,759	2,924
239	foro romano	0,00399	0,027	48,786	0,585
240	franquicia	0,00214	0,027	48,813	0,585
241	frutería	0,01303	0,135	48,948	2,924
242	fuelle	0,10379	0,836	49,784	18,129
243	galería	0,00886	0,054	49,838	1,170
244	galería de arte	0,01827	0,108	49,946	2,339
245	garaje	0,00314	0,027	49,973	0,585
246	gasolinera	0,01664	0,135	50,108	2,924
247	gato	0,01527	0,135	50,243	2,924
248	gaviota	0,00272	0,027	50,270	0,585
249	gente	0,15670	0,998	51,268	21,637
250	gerencia	0,00399	0,027	51,294	0,585
251	gestoría	0,00380	0,027	51,321	0,585
252	gimnasio	0,01613	0,135	51,456	2,924
253	glorieta	0,00551	0,081	51,537	1,754
254	golondrina	0,00285	0,027	51,564	0,585
255	gran superficie	0,00427	0,054	51,618	1,170
256	grandes almacenes	0,01151	0,081	51,699	1,754
257	grito	0,00380	0,027	51,726	0,585
258	grúa	0,00418	0,027	51,753	0,585
259	guardería	0,03365	0,324	52,077	7,018
260	guardia	0,00247	0,027	52,104	0,585
261	habitante	0,00690	0,054	52,157	1,170
262	hamaca	0,00214	0,027	52,184	0,585
263	heladería	0,00595	0,081	52,265	1,754
264	helipuerto	0,00362	0,027	52,292	0,585
265	hipermercado	0,00439	0,027	52,319	0,585
266	Hiperronda	0,00115	0,027	52,346	0,585
267	hogar	0,00194	0,027	52,373	0,585
268	hogar de ancianos	0,00177	0,027	52,400	0,585
269	hogar de jubilados	0,00285	0,027	52,427	0,585
270	hospital	0,15957	1,214	53,641	26,316
271	hostal	0,00938	0,108	53,749	2,339
272	hotel	0,02373	0,216	53,964	4,678
273	humo	0,01641	0,108	54,072	2,339
274	iglesia	0,14265	1,052	55,124	22,807
275	individuo	0,00557	0,027	55,151	0,585
276	inmobiliaria	0,00299	0,027	55,178	0,585
277	inmueble	0,00345	0,027	55,205	0,585
278	institución	0,00745	0,054	55,259	1,170

279	instituto	0,06665	0,566	55,825	12,281
280	instituto de enseñanza secundaria	0,00299	0,027	55,852	0,585
281	isleta	0,00235	0,027	55,879	0,585
282	jardín	0,11733	0,890	56,769	19,298
283	jardín de infancia	0,00272	0,027	56,796	0,585
284	jardinera	0,00235	0,027	56,823	0,585
285	jefatura de policía	0,01022	0,108	56,931	2,339
286	joven	0,00272	0,027	56,958	0,585
287	joyería	0,00314	0,027	56,985	0,585
288	juguetería	0,00664	0,054	57,039	1,170
289	Junta de Distrito	0,00259	0,027	57,066	0,585
290	juzgado	0,00663	0,054	57,120	1,170
291	lago	0,01162	0,081	57,201	1,754
292	lámpara	0,00204	0,027	57,228	0,585
293	lavandería	0,00285	0,027	57,255	0,585
294	librería	0,04578	0,378	57,632	8,187
295	línea de autobús	0,00613	0,054	57,686	1,170
296	luz	0,03898	0,297	57,983	6,433
297	Madrid	0,00329	0,027	58,010	0,585
298	Málaga	0,00299	0,027	58,037	0,585
299	manzana	0,01133	0,081	58,118	1,754
300	mar	0,02112	0,189	58,306	4,094
301	marquesina	0,00604	0,054	58,360	1,170
302	mascota	0,00585	0,054	58,414	1,170
303	McDonald's	0,00604	0,054	58,468	1,170
304	mediana	0,00224	0,027	58,495	0,585
305	medio de transporte	0,00314	0,027	58,522	0,585
306	mendigo	0,00168	0,027	58,549	0,585
307	mercadillo	0,02087	0,189	58,738	4,094
308	mercado	0,04939	0,378	59,115	8,187
309	mercería	0,00806	0,081	59,196	1,754
310	metro	0,13208	0,971	60,167	21,053
311	ministerio	0,00329	0,027	60,194	0,585
312	mirador	0,00380	0,054	60,248	1,170
313	moda	0,00380	0,027	60,275	0,585
314	monolito	0,00804	0,054	60,329	1,170
315	montaña	0,01281	0,108	60,437	2,339
316	monumento	0,11584	0,836	61,273	18,129
317	moto(cicleta)	0,08895	0,566	61,839	12,281
318	muchedumbre	0,00146	0,027	61,866	0,585
319	muelle	0,00506	0,027	61,893	0,585
320	museo	0,18254	1,268	63,161	27,485
321	nave industrial	0,00285	0,027	63,188	0,585
322	negocio	0,00801	0,054	63,242	1,170

323	niño	0,01895	0,189	63,430	4,094
324	notaría	0,00362	0,027	63,457	0,585
325	nube	0,00235	0,027	63,484	0,585
326	Nueva York	0,00153	0,027	63,511	0,585
327	obelisco	0,00995	0,081	63,592	1,754
328	obra	0,02717	0,243	63,835	5,263
329	ocio	0,00247	0,027	63,862	0,585
330	oficina	0,05358	0,405	64,266	8,772
331	oficina de empleo	0,01115	0,054	64,320	1,170
332	oficina del INEM	0,01846	0,108	64,428	2,339
333	ópera	0,01056	0,081	64,509	1,754
334	óptica	0,00385	0,054	64,563	1,170
335	orfanato	0,00132	0,027	64,590	0,585
336	organismo público	0,00329	0,027	64,617	0,585
337	outlet	0,00380	0,027	64,644	0,585
338	pabellón	0,00460	0,027	64,671	0,585
339	pájaro	0,00717	0,081	64,752	1,754
340	palacete	0,00247	0,027	64,779	0,585
341	palacio	0,02718	0,216	64,995	4,678
342	palacio de congresos	0,00779	0,054	65,049	1,170
343	panadería	0,09704	0,782	65,831	16,959
344	papelera	0,05903	0,512	66,343	11,111
345	papelería	0,00932	0,081	66,424	1,754
346	parada	0,01307	0,108	66,532	2,339
347	parada de autobús	0,04642	0,378	66,909	8,187
348	parada de taxi	0,00177	0,027	66,936	0,585
349	parking	0,02655	0,243	67,179	5,263
350	parque	0,38495	2,670	69,849	57,895
351	parque de bomberos	0,00966	0,081	69,930	1,754
352	pasaje	0,00778	0,054	69,984	1,170
353	pasarela	0,00168	0,027	70,011	0,585
354	paseo	0,03623	0,270	70,280	5,848
355	paseo marítimo	0,06372	0,485	70,766	10,526
356	paso de cebra	0,02659	0,189	70,955	4,094
357	paso de peatones	0,11108	0,809	71,764	17,544
358	pastelería	0,01172	0,108	71,872	2,339
359	patín	0,00362	0,027	71,899	0,585
360	patio	0,00329	0,027	71,926	0,585
361	peaje	0,00399	0,027	71,953	0,585
362	peatón	0,07855	0,512	72,465	11,111
363	pelea	0,00460	0,027	72,492	0,585
364	película	0,00399	0,027	72,519	0,585
365	peluquería	0,01841	0,162	72,681	3,509
366	perfumería	0,00314	0,027	72,708	0,585

367	periferia	0,00399	0,027	72,735	0,585
368	perrera	0,00418	0,027	72,762	0,585
369	perro	0,03524	0,324	73,085	7,018
370	persona	0,05300	0,351	73,436	7,602
371	pescadería	0,03518	0,297	73,732	6,433
372	piscina	0,01931	0,189	73,921	4,094
373	piscina municipal	0,00285	0,027	73,948	0,585
374	piso	0,09707	0,701	74,649	15,205
375	pito	0,00329	0,027	74,676	0,585
376	pizzería	0,00299	0,027	74,703	0,585
377	plano	0,00483	0,027	74,730	0,585
378	planta	0,01207	0,135	74,865	2,924
379	playa	0,07422	0,620	75,485	13,450
380	plaza	0,26690	1,807	77,292	39,181
381	plaza de toros	0,00929	0,081	77,373	1,754
382	Plaza Mayor	0,00204	0,027	77,400	0,585
383	policía	0,06311	0,485	77,886	10,526
384	polideportivo	0,03144	0,270	78,155	5,848
385	polígono	0,00991	0,081	78,236	1,754
386	polución	0,03659	0,270	78,506	5,848
387	pozo	0,00285	0,027	78,533	0,585
388	prisa	0,01267	0,081	78,614	1,754
389	prisión	0,00272	0,027	78,641	0,585
390	problema	0,00506	0,027	78,668	0,585
391	psiquiátrico	0,00139	0,027	78,695	0,585
392	pub	0,05591	0,485	79,180	10,526
393	pueblo	0,01304	0,081	79,261	1,754
394	puente	0,09148	0,728	79,989	15,789
395	puerto	0,07529	0,566	80,556	12,281
396	puesto	0,00285	0,027	80,583	0,585
397	puesto de flores	0,00194	0,027	80,609	0,585
398	quiosco	0,03361	0,270	80,879	5,848
399	radar	0,00362	0,027	80,906	0,585
400	rambla	0,00483	0,027	80,933	0,585
401	raqueta	0,00345	0,027	80,960	0,585
402	rascacielos	0,07251	0,566	81,526	12,281
403	rebajas	0,00418	0,027	81,553	0,585
404	redonda	0,00380	0,027	81,580	0,585
405	relojería	0,00628	0,054	81,634	1,170
406	RENFE	0,00684	0,054	81,688	1,170
407	resbaladera	0,00214	0,027	81,715	0,585
408	residencia de ancianos	0,00259	0,027	81,742	0,585
409	restaurante	0,19179	1,510	83,252	32,749
410	resto arqueológico	0,00194	0,027	83,279	0,585
411	río	0,03525	0,297	83,576	6,433

412	ronda	0,00299	0,027	83,603	0,585
413	rotonda	0,13413	0,998	84,601	21,637
414	ruido	0,08896	0,620	85,221	13,450
415	salir	0,00214	0,027	85,248	0,585
416	sede	0,00285	0,027	85,275	0,585
417	semáforo	0,40953	2,616	87,891	56,725
418	señal	0,03946	0,297	88,188	6,433
419	señal de tráfico	0,03606	0,324	88,511	7,018
420	señalización	0,00506	0,027	88,538	0,585
421	servicio	0,00439	0,027	88,565	0,585
422	seto	0,00224	0,027	88,592	0,585
423	sirena	0,00460	0,027	88,619	0,585
424	sombra	0,00194	0,027	88,646	0,585
425	sonido	0,00314	0,027	88,673	0,585
426	souvenir	0,00168	0,027	88,700	0,585
427	stop	0,00399	0,027	88,727	0,585
428	suburbio	0,00937	0,054	88,781	1,170
429	suciedad	0,00894	0,081	88,862	1,754
430	suelo	0,00475	0,054	88,916	1,170
431	supermercado	0,11842	0,944	89,860	20,468
432	taberna	0,00185	0,027	89,887	0,585
433	tanatorio	0,00204	0,027	89,914	0,585
434	tapería	0,00654	0,054	89,968	1,170
435	taxi	0,09244	0,701	90,669	15,205
436	teatro	0,14916	1,106	91,775	23,977
437	teatro romano	0,00717	0,054	91,828	1,170
438	tejado	0,00185	0,027	91,855	0,585
439	tensión	0,00177	0,027	91,882	0,585
440	terrazza	0,00677	0,054	91,936	1,170
441	tetería	0,01440	0,135	92,071	2,924
442	tiempo	0,00259	0,027	92,098	0,585
443	tienda	0,38147	2,751	94,849	59,649
444	tienda de ropa	0,00902	0,108	94,957	2,339
445	tienda de ultramarinos	0,00409	0,054	95,011	1,170
446	tierra	0,00506	0,027	95,038	0,585
447	tintorería	0,00247	0,027	95,065	0,585
448	tobogán	0,00471	0,054	95,119	1,170
449	torre	0,00604	0,054	95,173	1,170
450	trabajo	0,02546	0,216	95,388	4,678
451	tráfico	0,15439	0,998	96,386	21,637
452	transeúnte	0,01177	0,081	96,467	1,754
453	transporte	0,01376	0,081	96,548	1,754
454	transporte público	0,01259	0,108	96,656	2,339
455	tranvía	0,02653	0,243	96,899	5,263
456	tren	0,05925	0,458	97,357	9,942

457	túnel	0,01433	0,135	97,492	2,924
458	turismo	0,00418	0,027	97,519	0,585
459	turista	0,01145	0,108	97,627	2,339
460	universidad	0,13232	1,106	98,732	23,977
461	universidad privada	0,00362	0,027	98,759	0,585
462	universidad pública	0,00380	0,027	98,786	0,585
463	urbanización	0,02476	0,243	99,029	5,263
464	valle	0,00272	0,027	99,056	0,585
465	vecindario	0,00439	0,027	99,083	0,585
466	vecino	0,00439	0,027	99,110	0,585
467	vehículo	0,00905	0,054	99,164	1,170
468	ventana	0,00314	0,027	99,191	0,585
469	verja	0,00557	0,027	99,218	0,585
470	veterinario	0,00399	0,027	99,245	0,585
471	vía	0,00483	0,027	99,272	0,585
472	viandante	0,00665	0,054	99,326	1,170
473	videoclub	0,00214	0,027	99,353	0,585
474	viejo	0,00285	0,027	99,380	0,585
475	vivienda	0,01005	0,081	99,461	1,754
476	zapatería	0,01280	0,162	99,622	3,509
477	zona azul	0,00460	0,027	99,649	0,585
478	zona peatonal	0,00439	0,027	99,676	0,585
479	zona recreativa	0,00549	0,054	99,730	1,170
480	zona residencial	0,01217	0,081	99,811	1,754
481	zona verde	0,01014	0,054	99,865	1,170
482	zona vip	0,00259	0,027	99,892	0,585
483	zoológico	0,01277	0,108	100,000	2,339

1.1.6. Medios de transporte

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(auto)bus	0,65868	4,960	4,960	76,608
2	(auto)bus escolar	0,00194	0,038	4,998	0,585
3	(auto)bus turístico	0,00204	0,038	5,036	0,585
4	a pie	0,07413	0,757	5,793	11,696
5	aeroplano	0,06855	0,644	6,437	9,942
6	aeropuerto	0,00647	0,076	6,513	1,170
7	ala delta	0,00834	0,114	6,626	1,754
8	ambulancia	0,00719	0,076	6,702	1,170
9	andar	0,00549	0,076	6,778	1,170
10	animal	0,00345	0,038	6,816	0,585
11	aparcamiento	0,00139	0,038	6,853	0,585
12	ascensor	0,00652	0,076	6,929	1,170
13	asno	0,01122	0,114	7,043	1,754
14	AT-AT	0,00483	0,038	7,081	0,585
15	AT-ST	0,00460	0,038	7,119	0,585
16	autocar	0,06937	0,530	7,649	8,187
17	autocaravana	0,01081	0,114	7,762	1,754
18	automóvil	0,00782	0,076	7,838	1,170
19	autopista	0,00160	0,038	7,876	0,585
20	autovía	0,00168	0,038	7,914	0,585
21	AVE	0,05640	0,530	8,444	8,187
22	avestruz	0,00185	0,038	8,482	0,585
23	avión	0,72650	5,907	14,388	91,228
24	avión de caza	0,00362	0,038	14,426	0,585
25	avioneta	0,14597	1,515	15,941	23,392
26	balsa	0,00362	0,038	15,979	0,585
27	barca	0,10556	1,136	17,115	17,544
28	barco	0,59187	5,150	22,264	79,532
29	barco de vapor	0,00670	0,076	22,340	1,170
30	barco de vela	0,01008	0,114	22,454	1,754
31	berlina	0,00439	0,038	22,491	0,585
32	bici(cleta)	0,66119	5,377	27,868	83,041
33	bote	0,02713	0,303	28,171	4,678
34	buey	0,00946	0,114	28,285	1,754
35	buque	0,00798	0,076	28,360	1,170
36	burro	0,09935	1,022	29,383	15,789
37	burro-taxi	0,02117	0,189	29,572	2,924
38	caballo	0,23809	2,461	32,033	38,012
39	calesa	0,00926	0,114	32,147	1,754
40	camello	0,05136	0,606	32,753	9,357
41	caminar	0,00418	0,038	32,791	0,585
42	camión	0,27092	2,537	35,328	39,181
43	camión de bomberos	0,00439	0,038	35,365	0,585
44	camioneta	0,03255	0,341	35,706	5,263

45	canoa	0,04197	0,417	36,123	6,433
46	caravana	0,05434	0,530	36,653	8,187
47	carreta	0,01005	0,114	36,766	1,754
48	carretilla	0,00329	0,038	36,804	0,585
49	carricoche	0,00744	0,076	36,880	1,170
50	carro	0,02651	0,303	37,183	4,678
51	carro de caballos	0,00224	0,038	37,221	0,585
52	carromato	0,00966	0,114	37,334	1,754
53	carroza	0,01202	0,151	37,486	2,339
54	carruaje	0,01667	0,189	37,675	2,924
55	catamarán	0,00797	0,076	37,751	1,170
56	cayuco	0,00808	0,114	37,864	1,754
57	Cercanías	0,03649	0,341	38,205	5,263
58	ciclomotor	0,09063	0,795	39,000	12,281
59	coche	0,85702	5,983	44,983	92,398
60	coche de bebé	0,00285	0,038	45,021	0,585
61	coche de caballos	0,08806	0,947	45,967	14,620
62	coche deportivo	0,00763	0,076	46,043	1,170
63	coche descapotable	0,02425	0,227	46,270	3,509
64	coche particular	0,00966	0,076	46,346	1,170
65	coche teledirigido	0,00285	0,038	46,384	0,585
66	cohete	0,00247	0,038	46,422	0,585
67	cohete espacial	0,02877	0,265	46,687	4,094
68	crucero	0,11806	1,287	47,974	19,883
69	Cuatro Latas	0,00460	0,038	48,012	0,585
70	cuatro por cuatro	0,02831	0,227	48,239	3,509
71	diligencia	0,00823	0,114	48,353	1,754
72	dirigible	0,01602	0,189	48,542	2,924
73	dromedario	0,01249	0,151	48,694	2,339
74	elefante	0,01500	0,189	48,883	2,924
75	embarcación	0,00299	0,038	48,921	0,585
76	embarque	0,00214	0,038	48,959	0,585
77	equipaje	0,00204	0,038	48,997	0,585
78	escalera mecánica	0,00564	0,076	49,072	1,170
79	esquí	0,01751	0,227	49,300	3,509
80	esquí acuático	0,00299	0,038	49,337	0,585
81	estación de autobús	0,00329	0,038	49,375	0,585
82	estación de metro	0,00299	0,038	49,413	0,585
83	estación de tren	0,00314	0,038	49,451	0,585
84	excavadora	0,00259	0,038	49,489	0,585
85	facturar	0,00194	0,038	49,527	0,585
86	ferrocarril	0,01315	0,151	49,678	2,339
87	ferry	0,06557	0,682	50,360	10,526
88	funicular	0,01994	0,189	50,549	2,924
89	furgón	0,01162	0,114	50,663	1,754
90	furgoneta	0,19531	1,855	52,518	28,655
91	globo	0,08723	0,909	53,427	14,035
92	globo aerostático	0,02482	0,265	53,692	4,094
93	góndola	0,01443	0,189	53,881	2,924

94	helicóptero	0,30242	2,991	56,872	46,199
95	helipuerto	0,00272	0,038	56,910	0,585
96	hidroavión	0,01206	0,114	57,024	1,754
97	hidropedal	0,02007	0,227	57,251	3,509
98	jet	0,00314	0,038	57,289	0,585
99	jet privado	0,01039	0,114	57,402	1,754
100	kayak	0,02557	0,265	57,668	4,094
101	kilómetro	0,00235	0,038	57,705	0,585
102	lancha	0,07419	0,757	58,463	11,696
103	limusina	0,01361	0,151	58,614	2,339
104	locomotora	0,00460	0,038	58,652	0,585
105	lomo de burro	0,00314	0,038	58,690	0,585
106	máquina del tiempo	0,00272	0,038	58,728	0,585
107	metro	0,47767	4,241	62,969	65,497
108	metro ligero	0,00881	0,076	63,044	1,170
109	microbús	0,00968	0,114	63,158	1,754
110	Mini	0,01667	0,151	63,309	2,339
111	minibús	0,00943	0,114	63,423	1,754
112	monociclo	0,00247	0,038	63,461	0,585
113	monopatín	0,18743	1,855	65,316	28,655
114	monorraíl	0,00710	0,076	65,392	1,170
115	monovolumen	0,00857	0,076	65,468	1,170
116	montacargas	0,00710	0,076	65,543	1,170
117	moto acuática	0,01874	0,227	65,771	3,509
118	moto de agua	0,00285	0,038	65,808	0,585
119	moto(cicleta)	0,59917	4,922	70,731	76,023
120	mula	0,00573	0,076	70,807	1,170
121	mulo	0,02274	0,227	71,034	3,509
122	nave espacial	0,02544	0,265	71,299	4,094
123	noria	0,00329	0,038	71,337	0,585
124	ómnibus	0,00676	0,076	71,412	1,170
125	paracaídas	0,01734	0,189	71,602	2,924
126	parada de autobús	0,00247	0,038	71,640	0,585
127	parapente	0,02982	0,303	71,942	4,678
128	parking	0,00132	0,038	71,980	0,585
129	patera	0,06873	0,757	72,738	11,696
130	patín	0,27118	2,688	75,426	41,520
131	patineta	0,01862	0,189	75,615	2,924
132	patinete	0,12943	1,363	76,978	21,053
133	peaje	0,00153	0,038	77,016	0,585
134	peatón	0,00146	0,038	77,054	0,585
135	pie	0,02776	0,303	77,357	4,678
136	pierna	0,00285	0,038	77,395	0,585
137	piragua	0,00816	0,114	77,509	1,754
138	planeador	0,00439	0,038	77,546	0,585
139	poni	0,00314	0,038	77,584	0,585
140	portaaviones	0,00548	0,076	77,660	1,170
141	potro	0,00768	0,076	77,736	1,170
142	puerto	0,00345	0,038	77,774	0,585

143	quad	0,04282	0,492	78,266	7,602
144	ranchera	0,00732	0,076	78,342	1,170
145	remolque	0,00507	0,076	78,417	1,170
146	Renfe	0,00285	0,038	78,455	0,585
147	reno	0,00712	0,076	78,531	1,170
148	scooter	0,02311	0,189	78,720	2,924
149	Seiscientos	0,01267	0,114	78,834	1,754
150	sidecar	0,04271	0,492	79,326	7,602
151	silla de ruedas	0,00299	0,038	79,364	0,585
152	skate	0,00546	0,076	79,440	1,170
153	submarino	0,10256	1,060	80,500	16,374
154	tabla de snowboard	0,00813	0,114	80,613	1,754
155	tabla de surf	0,00962	0,114	80,727	1,754
156	tándem	0,01303	0,151	80,878	2,339
157	taxi	0,33026	2,764	83,643	42,690
158	telecabina	0,00494	0,076	83,718	1,170
159	teleférico	0,02804	0,303	84,021	4,678
160	telesilla	0,00599	0,076	84,097	1,170
161	todoterreno	0,02827	0,265	84,362	4,094
162	tractor	0,04388	0,530	84,892	8,187
163	tráfico	0,00177	0,038	84,930	0,585
164	tráiler	0,06407	0,568	85,498	8,772
165	transatlántico	0,05990	0,682	86,179	10,526
166	transbordador	0,01300	0,114	86,293	1,754
167	transmediterráneo	0,00272	0,038	86,331	0,585
168	transporte público	0,00506	0,038	86,369	0,585
169	tranvía	0,32729	2,953	89,322	45,614
170	tren	0,70709	5,680	95,002	87,719
171	tren de alta velocidad	0,01714	0,151	95,153	2,339
172	tren de mercancías	0,00362	0,038	95,191	0,585
173	tren eléctrico	0,00329	0,038	95,229	0,585
174	triciclo	0,06225	0,719	95,949	11,111
175	trineo	0,03235	0,379	96,327	5,848
176	trolebús	0,03975	0,341	96,668	5,263
177	tuk-tuk	0,00483	0,038	96,706	0,585
178	ultraligero	0,00418	0,038	96,744	0,585
179	utilitario	0,00329	0,038	96,782	0,585
180	vaporetto	0,00506	0,038	96,819	0,585
181	vehículo	0,01038	0,076	96,895	1,170
182	velero	0,07116	0,757	97,652	11,696
183	velocidad	0,00185	0,038	97,690	0,585
184	vespa	0,00362	0,038	97,728	0,585
185	viaje	0,00224	0,038	97,766	0,585
186	yate	0,19211	1,893	99,659	29,240
187	yegua	0,00674	0,076	99,735	1,170
188	zapato	0,00585	0,038	99,773	0,585
189	zepelín	0,01670	0,189	99,962	2,924
190	zodiaco	0,00418	0,038	100,000	0,585

1.1.7. Juegos y distracciones

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(juego de) cartas	0,31110	2,300	2,300	40,936
2	(máquina) tragaperras	0,00592	0,066	2,365	1,170
3	¿Dónde está Wally?	0,00439	0,033	2,398	0,585
4	¿Quién es quién?	0,00362	0,033	2,431	0,585
5	acampar	0,00506	0,033	2,464	0,585
6	aerobic	0,01885	0,131	2,595	2,339
7	ahorcado	0,00745	0,066	2,661	1,170
8	ajedrez	0,26909	1,971	4,632	35,088
9	alcohol	0,00531	0,033	4,665	0,585
10	alpinismo	0,01322	0,131	4,796	2,339
11	amigo	0,00951	0,099	4,895	1,754
12	andar	0,00718	0,066	4,961	1,170
13	animal	0,00272	0,033	4,993	0,585
14	apostar	0,00272	0,033	5,026	0,585
15	aprender	0,00259	0,033	5,059	0,585
16	apuesta	0,00362	0,033	5,092	0,585
17	aquagym	0,00285	0,033	5,125	0,585
18	aro	0,00259	0,033	5,158	0,585
19	arte	0,00911	0,066	5,223	1,170
20	astronomía	0,00380	0,033	5,256	0,585
21	atletismo	0,07331	0,657	5,913	11,696
22	automovilismo	0,00362	0,033	5,946	0,585
23	ayudar	0,00259	0,033	5,979	0,585
24	bádminton	0,02415	0,230	6,209	4,094
25	bailar	0,15653	1,314	7,523	23,392
26	baile	0,03532	0,296	7,819	5,263
27	balconing	0,00259	0,033	7,852	0,585
28	ballet	0,00615	0,066	7,917	1,170
29	balón	0,01633	0,131	8,049	2,339
30	balón prisionero	0,00272	0,033	8,081	0,585
31	baloncesto	0,36446	2,792	10,874	49,708
32	balonmano	0,11528	0,986	11,859	17,544
33	balonvolea	0,00677	0,066	11,925	1,170
34	bar	0,04028	0,329	12,254	5,848
35	baraja	0,00418	0,033	12,286	0,585
36	Barbie	0,00957	0,099	12,385	1,754
37	barco	0,00789	0,066	12,451	1,170
38	beber	0,01763	0,164	12,615	2,924
39	béisbol	0,02161	0,197	12,812	3,509
40	bicicleta	0,07540	0,690	13,502	12,281
41	bicicleta elíptica	0,00399	0,033	13,535	0,585
42	bicicleta estática	0,00380	0,033	13,568	0,585

43	billar	0,03404	0,263	13,830	4,678
44	bingo	0,02971	0,263	14,093	4,678
45	blog	0,00299	0,033	14,126	0,585
46	bolera	0,00362	0,033	14,159	0,585
47	bolo	0,06019	0,493	14,652	8,772
48	bordar	0,00418	0,033	14,685	0,585
49	botellón	0,00439	0,033	14,717	0,585
50	boxeo	0,00399	0,033	14,750	0,585
51	bricolaje	0,00418	0,033	14,783	0,585
52	brisca	0,05074	0,361	15,145	6,433
53	Bubble Bobble	0,00247	0,033	15,177	0,585
54	bucear	0,00999	0,099	15,276	1,754
55	buceo	0,00399	0,033	15,309	0,585
56	burro	0,00399	0,033	15,342	0,585
57	Buscaminas	0,00259	0,033	15,375	0,585
58	Buzz!	0,00557	0,033	15,407	0,585
59	cabalgar	0,00285	0,033	15,440	0,585
60	caballería	0,00259	0,033	15,473	0,585
61	caballo	0,02315	0,197	15,670	3,509
62	cafetería	0,00399	0,033	15,703	0,585
63	caminar	0,03140	0,329	16,032	5,848
64	camping	0,00285	0,033	16,064	0,585
65	campo	0,01351	0,131	16,196	2,339
66	canasta	0,00483	0,033	16,229	0,585
67	canasta revuelta	0,00177	0,033	16,261	0,585
68	canicas	0,01544	0,131	16,393	2,339
69	canoa	0,00247	0,033	16,426	0,585
70	cantar	0,07539	0,591	17,017	10,526
71	carambola	0,00506	0,033	17,050	0,585
72	carrera	0,00782	0,066	17,116	1,170
73	carrera de caballos	0,00272	0,033	17,148	0,585
74	carrera de coches	0,00842	0,099	17,247	1,754
75	carrera de sacos	0,00194	0,033	17,280	0,585
76	carricoche	0,00506	0,033	17,313	0,585
77	casitas	0,00285	0,033	17,346	0,585
78	catedral	0,00399	0,033	17,378	0,585
79	caza	0,01056	0,099	17,477	1,754
80	cazar	0,00299	0,033	17,510	0,585
81	CD	0,00299	0,033	17,543	0,585
82	cenar	0,00877	0,066	17,608	1,170
83	chapas	0,00235	0,033	17,641	0,585
84	chaquete	0,01469	0,099	17,740	1,754
85	charlar	0,02189	0,164	17,904	2,924
86	chat	0,01587	0,164	18,068	2,924
87	chatear	0,03824	0,394	18,463	7,018
88	chinchón	0,00956	0,066	18,528	1,170
89	churro pico o terna	0,00314	0,033	18,561	0,585
90	ciclismo	0,05270	0,460	19,021	8,187
91	cine	0,35542	2,825	21,846	50,292

92	cinquillo	0,04969	0,394	22,240	7,018
93	circo	0,01044	0,099	22,339	1,754
94	clase	0,00299	0,033	22,372	0,585
95	Cluedo	0,03486	0,263	22,635	4,678
96	coche	0,00837	0,066	22,700	1,170
97	coche teledirigido	0,00418	0,033	22,733	0,585
98	cocinar	0,04308	0,427	23,160	7,602
99	coleccionar	0,00460	0,033	23,193	0,585
100	coleccionismo	0,00439	0,033	23,226	0,585
101	colorear	0,00483	0,033	23,259	0,585
102	columpio	0,01277	0,099	23,357	1,754
103	comba	0,07718	0,624	23,982	11,111
104	comecocos	0,00247	0,033	24,014	0,585
105	comer	0,01296	0,131	24,146	2,339
106	cometa	0,00399	0,033	24,179	0,585
107	cómic	0,00949	0,099	24,277	1,754
108	competición	0,00345	0,033	24,310	0,585
109	componer música	0,00399	0,033	24,343	0,585
110	compra	0,00299	0,033	24,376	0,585
111	comprar	0,00730	0,099	24,474	1,754
112	concierto	0,07018	0,624	25,099	11,111
113	conducir	0,00557	0,066	25,164	1,170
114	conejo de la suerte	0,00471	0,066	25,230	1,170
115	consola	0,02670	0,230	25,460	4,094
116	conversar	0,00224	0,033	25,493	0,585
117	correr	0,11668	1,051	26,544	18,713
118	corro de la patata	0,00531	0,066	26,610	1,170
119	coser	0,01492	0,131	26,741	2,339
120	crear	0,00362	0,033	26,774	0,585
121	cricket	0,00900	0,099	26,873	1,754
122	crucigrama	0,03574	0,296	27,168	5,263
123	cuatro en raya	0,00299	0,033	27,201	0,585
124	cuento	0,00859	0,066	27,267	1,170
125	culebrón	0,00235	0,033	27,300	0,585
126	dado	0,04954	0,394	27,694	7,018
127	damas	0,12673	1,051	28,745	18,713
128	danza	0,00728	0,066	28,811	1,170
129	dardo	0,01791	0,164	28,975	2,924
130	decoración	0,00460	0,033	29,008	0,585
131	decorar	0,00609	0,066	29,074	1,170
132	deporte	0,21584	1,708	30,782	30,409
133	deporte de aventura	0,00247	0,033	30,815	0,585
134	deporte de contacto	0,00285	0,033	30,848	0,585
135	deporte de riesgo	0,01197	0,131	30,979	2,339
136	descenso	0,00345	0,033	31,012	0,585
137	diario	0,00247	0,033	31,045	0,585
138	dibujar	0,02620	0,197	31,242	3,509
139	dibujo	0,00634	0,066	31,307	1,170
140	dibujo animado	0,00460	0,033	31,340	0,585

141	Dicciopinta	0,00314	0,033	31,373	0,585
142	discoteca	0,04943	0,427	31,800	7,602
143	disfraz	0,00272	0,033	31,833	0,585
144	divertirse	0,00259	0,033	31,866	0,585
145	documental	0,01834	0,164	32,030	2,924
146	dominó	0,05902	0,493	32,523	8,772
147	dormir	0,00345	0,033	32,556	0,585
148	DVD	0,00694	0,066	32,622	1,170
149	El Gordo	0,00380	0,033	32,654	0,585
150	el patio de mi casa	0,00557	0,033	32,687	0,585
151	elástico	0,05532	0,427	33,114	7,602
152	emborracharse	0,00445	0,066	33,180	1,170
153	equitación	0,02103	0,197	33,377	3,509
154	escalada	0,03082	0,296	33,673	5,263
155	escalar	0,00760	0,066	33,739	1,170
156	escalerilla	0,00460	0,033	33,771	0,585
157	escondite	0,14053	1,150	34,921	20,468
158	escribir	0,06546	0,591	35,512	10,526
159	escritura	0,01729	0,131	35,644	2,339
160	escuchar música	0,08758	0,723	36,367	12,865
161	esculpir	0,00439	0,033	36,399	0,585
162	escultura	0,01373	0,131	36,531	2,339
163	esgrima	0,00514	0,066	36,597	1,170
164	eslalon	0,00399	0,033	36,629	0,585
165	espectáculo	0,01573	0,131	36,761	2,339
166	espeleología	0,00314	0,033	36,794	0,585
167	esquí	0,01573	0,131	36,925	2,339
168	esquí acuático	0,00194	0,033	36,958	0,585
169	esquí alpino	0,00418	0,033	36,991	0,585
170	esquí de fondo	0,00439	0,033	37,024	0,585
171	esquiar	0,05616	0,493	37,516	8,772
172	estiramiento	0,00345	0,033	37,549	0,585
173	estudiar	0,00880	0,099	37,648	1,754
174	excursión	0,01265	0,131	37,779	2,339
175	excursión a la naturaleza	0,00235	0,033	37,812	0,585
176	explorar	0,00418	0,033	37,845	0,585
177	exposición	0,00224	0,033	37,878	0,585
178	Facebook	0,04584	0,394	38,272	7,018
179	Farm Ville	0,01014	0,066	38,338	1,170
180	feria	0,00314	0,033	38,371	0,585
181	ficción	0,00399	0,033	38,403	0,585
182	fitness	0,00506	0,033	38,436	0,585
183	flamenco	0,00259	0,033	38,469	0,585
184	flauta	0,00744	0,066	38,535	1,170
185	footing	0,01454	0,131	38,666	2,339
186	foro	0,00285	0,033	38,699	0,585
187	foto(grafía)	0,03060	0,263	38,962	4,678
188	freaky	0,00483	0,033	38,995	0,585

189	fumar porros	0,00153	0,033	39,028	0,585
190	funky	0,00272	0,033	39,060	0,585
191	fútbol	0,47019	3,384	42,444	60,234
192	fútbol americano	0,00285	0,033	42,477	0,585
193	fútbol playa	0,00259	0,033	42,510	0,585
194	fútbol sala	0,00439	0,033	42,543	0,585
195	futbolín	0,01261	0,099	42,641	1,754
196	galería de arte	0,00439	0,033	42,674	0,585
197	gallinita ciega	0,03374	0,296	42,970	5,263
198	Game Boy	0,00922	0,131	43,101	2,339
199	gastronomía	0,00329	0,033	43,134	0,585
200	gesto	0,00224	0,033	43,167	0,585
201	gimnasia	0,01345	0,131	43,298	2,339
202	gimnasia rítmica	0,00732	0,066	43,364	1,170
203	gimnasio	0,03442	0,329	43,693	5,848
204	golf	0,04797	0,394	44,087	7,018
205	Google	0,00345	0,033	44,120	0,585
206	guitarra	0,00819	0,066	44,185	1,170
207	hablar	0,01683	0,164	44,350	2,924
208	hablar por teléfono	0,00543	0,066	44,415	1,170
209	hablar por WhatsApp	0,00204	0,033	44,448	0,585
210	hacer cine	0,00399	0,033	44,481	0,585
211	halterofilia	0,00380	0,033	44,514	0,585
212	hermano	0,00399	0,033	44,547	0,585
213	hijo puta	0,00483	0,033	44,580	0,585
214	hípica	0,00177	0,033	44,612	0,585
215	hockey	0,02656	0,230	44,842	4,094
216	Hotmail	0,00709	0,066	44,908	1,170
217	idioma	0,00235	0,033	44,941	0,585
218	iglesia	0,00380	0,033	44,974	0,585
219	Ikea	0,00329	0,033	45,007	0,585
220	informativo	0,00418	0,033	45,039	0,585
221	intercambio de idiomas	0,00272	0,033	45,072	0,585
222	internet	0,15533	1,216	46,288	21,637
223	investigar	0,00399	0,033	46,321	0,585
224	ir de compras	0,03267	0,296	46,616	5,263
225	ir de tiendas	0,00768	0,066	46,682	1,170
226	jeroglífico	0,00483	0,033	46,715	0,585
227	jogging	0,00299	0,033	46,748	0,585
228	juego	0,00483	0,033	46,781	0,585
229	juego de azar	0,02444	0,230	47,011	4,094
230	juego de magia	0,00418	0,033	47,043	0,585
231	juego de mesa	0,14266	1,084	48,127	19,298
232	juego de naipes	0,00992	0,066	48,193	1,170
233	juego de ordenador	0,03818	0,263	48,456	4,678
234	juego de rol	0,00506	0,033	48,489	0,585
235	juego del chupito	0,00314	0,033	48,522	0,585

236	juego didáctico	0,00399	0,033	48,555	0,585
237	juego on line	0,00506	0,033	48,587	0,585
238	jugar	0,01162	0,099	48,686	1,754
239	Jungle Speed	0,00712	0,066	48,752	1,170
240	karaoke	0,05262	0,460	49,212	8,187
241	kárate	0,01418	0,131	49,343	2,339
242	kayak	0,00235	0,033	49,376	0,585
243	kickboxing	0,00473	0,066	49,442	1,170
244	lectura	0,07517	0,591	50,033	10,526
245	leer	0,22765	1,840	51,873	32,749
246	libro	0,05449	0,427	52,300	7,602
247	limpiar la casa	0,00380	0,033	52,332	0,585
248	lotería	0,01799	0,164	52,497	2,924
249	mando a distancia	0,00380	0,033	52,530	0,585
250	maratón	0,00224	0,033	52,562	0,585
251	mascota	0,00576	0,066	52,628	1,170
252	máster	0,00585	0,033	52,661	0,585
253	masturbar	0,00362	0,033	52,694	0,585
254	médicos	0,00345	0,033	52,727	0,585
255	meditación	0,00235	0,033	52,760	0,585
256	mentiroso	0,02645	0,197	52,957	3,509
257	Messenger	0,02054	0,164	53,121	2,924
258	minigolf	0,00362	0,033	53,154	0,585
259	monólogo	0,00299	0,033	53,187	0,585
260	monopatín	0,00527	0,066	53,252	1,170
261	Monopoly	0,12273	0,986	54,238	17,544
262	montañismo	0,00931	0,099	54,336	1,754
263	monumento	0,00754	0,066	54,402	1,170
264	motociclismo	0,02397	0,197	54,599	3,509
265	móvil	0,00644	0,099	54,698	1,754
266	muñeco	0,03147	0,296	54,993	5,263
267	mus	0,07108	0,558	55,552	9,942
268	musculación	0,00345	0,033	55,585	0,585
269	museo	0,05192	0,460	56,045	8,187
270	música	0,20842	1,675	57,720	29,825
271	musical	0,01946	0,164	57,884	2,924
272	nadar	0,03931	0,361	58,246	6,433
273	Nancy	0,00247	0,033	58,279	0,585
274	natación	0,12190	1,051	59,330	18,713
275	naturaleza	0,00531	0,033	59,363	0,585
276	navegar	0,01389	0,164	59,527	2,924
277	Nintendo	0,00738	0,066	59,593	1,170
278	noria	0,00439	0,033	59,625	0,585
279	noticia	0,00506	0,033	59,658	0,585
280	novela	0,01561	0,164	59,823	2,924
281	obra	0,00801	0,066	59,888	1,170

282	oca	0,24883	1,840	61,728	32,749
283	ocio nocturno	0,00299	0,033	61,761	0,585
284	ola	0,00329	0,033	61,794	0,585
285	ópera	0,01185	0,099	61,892	1,754
286	ordenador	0,07938	0,657	62,549	11,696
287	pádel	0,13333	1,150	63,699	20,468
288	paintball	0,00235	0,033	63,732	0,585
289	palacio	0,00418	0,033	63,765	0,585
290	paleta	0,00612	0,066	63,830	1,170
291	palmas	0,00863	0,099	63,929	1,754
292	palomita	0,00585	0,033	63,962	0,585
293	paracaídas	0,00224	0,033	63,995	0,585
294	parchís	0,44286	3,055	67,050	54,386
295	pares o nones	0,00659	0,066	67,116	1,170
296	parque	0,00380	0,033	67,148	0,585
297	parque de atracciones	0,01151	0,099	67,247	1,754
298	partido	0,00557	0,033	67,280	0,585
299	Party & Co.	0,02492	0,230	67,510	4,094
300	pasear	0,15688	1,380	68,890	24,561
301	pasear al perro	0,00630	0,066	68,955	1,170
302	pasear animales de compañía	0,00194	0,033	68,988	0,585
303	pasear por el campo	0,00285	0,033	69,021	0,585
304	paseo	0,00742	0,066	69,087	1,170
305	patín	0,00768	0,066	69,152	1,170
306	patinaje	0,01181	0,099	69,251	1,754
307	patinaje sobre hielo	0,00224	0,033	69,284	0,585
308	patinar	0,07681	0,690	69,974	12,281
309	patinete	0,00460	0,033	70,007	0,585
310	patio de mi casa	0,00153	0,033	70,039	0,585
311	película	0,06992	0,591	70,631	10,526
312	pelota	0,04340	0,296	70,926	5,263
313	pelota vasca	0,00224	0,033	70,959	0,585
314	peluquería	0,00362	0,033	70,992	0,585
315	pensar	0,00531	0,066	71,058	1,170
316	peonza	0,00259	0,033	71,091	0,585
317	periódico	0,00432	0,066	71,156	1,170
318	pesas	0,00235	0,033	71,189	0,585
319	pesca	0,01027	0,099	71,288	1,754
320	pescar	0,00819	0,066	71,353	1,170
321	Pet Society	0,00506	0,033	71,386	0,585
322	petanca	0,07439	0,591	71,978	10,526
323	piano	0,00362	0,033	72,011	0,585
324	Pictionary	0,01375	0,099	72,109	1,754
325	piedra papel o tijera	0,00691	0,066	72,175	1,170

326	Pilates	0,01003	0,099	72,273	1,754
327	pillapilla	0,09208	0,788	73,062	14,035
328	ping-pong	0,02637	0,263	73,325	4,678
329	pintar	0,06285	0,526	73,850	9,357
330	pintura	0,02920	0,230	74,080	4,094
331	piragüismo	0,00677	0,066	74,146	1,170
332	piso	0,01125	0,099	74,244	1,754
333	Play Station	0,07299	0,624	74,869	11,111
334	playa	0,04356	0,394	75,263	7,018
335	Playmobil	0,00781	0,099	75,361	1,754
336	Pokemon	0,00272	0,033	75,394	0,585
337	policías y ladrones	0,00906	0,099	75,493	1,754
338	poli-ladro	0,00460	0,033	75,526	0,585
339	pollito inglés	0,00985	0,099	75,624	1,754
340	póquer	0,12705	0,953	76,577	16,959
341	Pro Evolution Soccer	0,00840	0,066	76,643	1,170
342	programa	0,00285	0,033	76,675	0,585
343	psicólogo	0,00185	0,033	76,708	0,585
344	PSP	0,00628	0,066	76,774	1,170
345	pub	0,01800	0,164	76,938	2,924
346	puenting	0,01145	0,131	77,070	2,339
347	pulso	0,00329	0,033	77,102	0,585
348	pulso chino	0,00314	0,033	77,135	0,585
349	punto de cruz	0,00571	0,066	77,201	1,170
350	puzle	0,02360	0,197	77,398	3,509
351	quiniela	0,00380	0,033	77,431	0,585
352	radio	0,03991	0,361	77,792	6,433
353	rafting	0,00345	0,033	77,825	0,585
354	rayuela	0,02368	0,197	78,022	3,509
355	recortable	0,00285	0,033	78,055	0,585
356	red social	0,02503	0,164	78,219	2,924
357	reguetón	0,00247	0,033	78,252	0,585
358	relevo	0,00735	0,066	78,318	1,170
359	restaurante	0,02538	0,197	78,515	3,509
360	reunirse con amigos	0,00299	0,033	78,548	0,585
361	revista	0,02169	0,197	78,745	3,509
362	rey del tres	0,00439	0,033	78,778	0,585
363	Risk	0,01220	0,099	78,876	1,754
364	rocódromo	0,00483	0,033	78,909	0,585
365	rol	0,00557	0,033	78,942	0,585
366	ronda	0,00399	0,033	78,975	0,585
367	rugby	0,04982	0,427	79,402	7,602
368	ruleta	0,00639	0,066	79,468	1,170
369	ruleta rusa	0,00483	0,033	79,501	0,585
370	Rummikub	0,00585	0,033	79,534	0,585

371	salir	0,06151	0,526	80,059	9,357
372	salir a cenar	0,01641	0,164	80,223	2,924
373	salir con amigos	0,10830	0,887	81,110	15,789
374	salir de fiesta	0,01535	0,164	81,275	2,924
375	salir de marcha	0,01136	0,099	81,373	1,754
376	salsa	0,00703	0,066	81,439	1,170
377	saltar	0,01511	0,131	81,570	2,339
378	salto de altura	0,00299	0,033	81,603	0,585
379	salto de la burra	0,00329	0,033	81,636	0,585
380	salto de longitud	0,00299	0,033	81,669	0,585
381	salto de pértiga	0,00771	0,066	81,735	1,170
382	sauna	0,00247	0,033	81,767	0,585
383	Scattergories	0,02048	0,164	81,932	2,924
384	Scrabble	0,03619	0,296	82,227	5,263
385	Sega Mega Drive	0,00224	0,033	82,260	0,585
386	senderismo	0,08258	0,756	83,016	13,450
387	serie	0,04599	0,361	83,377	6,433
388	serie de televisión	0,00418	0,033	83,410	0,585
389	Series Yonkis	0,01084	0,099	83,509	1,754
390	sexo	0,00483	0,033	83,541	0,585
391	siesta	0,00857	0,066	83,607	1,170
392	siete diferencias	0,00460	0,033	83,640	0,585
393	Sims	0,00460	0,033	83,673	0,585
394	SingStar	0,01329	0,099	83,771	1,754
395	snow(board)	0,01421	0,131	83,903	2,339
396	soldado	0,00285	0,033	83,936	0,585
397	solitario	0,02022	0,164	84,100	2,924
398	Sonic	0,00314	0,033	84,133	0,585
399	sopa de letras	0,00832	0,066	84,198	1,170
400	spinning	0,02074	0,164	84,363	2,924
401	squash	0,00905	0,066	84,428	1,170
402	step	0,01794	0,131	84,560	2,339
403	subir al monte	0,00204	0,033	84,593	0,585
404	submarinismo	0,00658	0,066	84,658	1,170
405	sudoku	0,02688	0,230	84,888	4,094
406	Super Mario	0,01571	0,164	85,053	2,924
407	Super Pang	0,00259	0,033	85,085	0,585
408	surf	0,00362	0,033	85,118	0,585
409	surfear	0,00599	0,066	85,184	1,170
410	taba	0,00272	0,033	85,217	0,585
411	Tabú	0,05422	0,526	85,742	9,357
412	taekwondo	0,00214	0,033	85,775	0,585
413	taichí	0,00345	0,033	85,808	0,585
414	tambor	0,00380	0,033	85,841	0,585
415	tapersex	0,00314	0,033	85,874	0,585
416	teatro	0,18386	1,544	87,418	27,485

417	tejer	0,00314	0,033	87,451	0,585
418	telenovela	0,00658	0,066	87,516	1,170
419	televisión	0,22784	1,807	89,323	32,164
420	tenis	0,24930	2,004	91,327	35,673
421	tenis de mesa	0,00235	0,033	91,360	0,585
422	tertulia	0,00380	0,033	91,393	0,585
423	test	0,00614	0,066	91,459	1,170
424	Tetris	0,02353	0,164	91,623	2,924
425	tienda	0,00285	0,033	91,656	0,585
426	tiovivo	0,00557	0,033	91,689	0,585
427	tirar piedras al tren	0,00380	0,033	91,721	0,585
428	tiro con arco	0,00725	0,066	91,787	1,170
429	tobogán	0,00399	0,033	91,820	0,585
430	tocar el piano	0,01224	0,099	91,919	1,754
431	tocar la guitarra	0,01877	0,131	92,050	2,339
432	tocar un instrumento	0,03919	0,296	92,346	5,263
433	tomar café	0,01169	0,099	92,444	1,754
434	tomar el sol	0,00859	0,066	92,510	1,170
435	toros	0,00557	0,033	92,543	0,585
436	trekking	0,00531	0,033	92,576	0,585
437	tres en línea	0,00557	0,033	92,608	0,585
438	tres en raya	0,03228	0,296	92,904	5,263
439	triatlón	0,01056	0,099	93,003	1,754
440	trineo	0,00460	0,033	93,035	0,585
441	Trivial	0,11859	0,854	93,890	15,205
442	trompo	0,00272	0,033	93,922	0,585
443	Tuenti	0,02719	0,197	94,120	3,509
444	turismo	0,00687	0,099	94,218	1,754
445	tute	0,01785	0,131	94,350	2,339
446	Twitter	0,01435	0,131	94,481	2,339
447	UNO	0,01095	0,131	94,612	2,339
448	vela	0,01051	0,099	94,711	1,754
449	veo-veo	0,01582	0,164	94,875	2,924
450	viajar	0,10444	0,887	95,762	15,789
451	viaje	0,01144	0,099	95,861	1,754
452	vídeo	0,00946	0,099	95,959	1,754
453	videoconsola	0,03767	0,263	96,222	4,678
454	videojuego	0,27037	2,070	98,292	36,842
455	violín	0,00418	0,033	98,325	0,585
456	voleibol	0,06613	0,591	98,916	10,526
457	voluntariado	0,00247	0,033	98,949	0,585
458	waterpolo	0,00736	0,099	99,047	1,754
459	Wii	0,03433	0,296	99,343	5,263
460	Wikipedia	0,00362	0,033	99,376	0,585
461	World of Warcraft	0,00362	0,033	99,409	0,585

462	Xbox	0,00686	0,066	99,474	1,170
463	Yathzee	0,00214	0,033	99,507	0,585
464	yoga	0,01646	0,164	99,671	2,924
465	Youtube	0,01230	0,099	99,770	1,754
466	yoyó	0,00557	0,033	99,803	0,585
467	yudo	0,01242	0,099	99,901	1,754
468	zapatilla por detrás	0,00224	0,033	99,934	0,585
469	Zelda	0,00768	0,066	100,000	1,170

1.1.8. Profesiones y oficios

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	abogado	0,29363	2,000	2,000	42,105
2	acomodador	0,00732	0,083	2,083	1,754
3	actor	0,14190	1,083	3,167	22,807
4	actor porno	0,00531	0,028	3,194	0,585
5	actriz	0,01850	0,111	3,306	2,339
6	administrador	0,00664	0,056	3,361	1,170
7	administrador de empresas	0,00362	0,028	3,389	0,585
8	administrativo	0,11288	0,861	4,250	18,129
9	aeronauta	0,00285	0,028	4,278	0,585
10	afilador	0,00652	0,056	4,333	1,170
11	agente	0,00314	0,028	4,361	0,585
12	agente de seguros	0,00314	0,028	4,389	0,585
13	agente de viajes	0,01202	0,083	4,472	1,754
14	agente inmobiliario	0,00584	0,056	4,528	1,170
15	agente secreto	0,00557	0,028	4,556	0,585
16	agricultor	0,08211	0,667	5,222	14,035
17	aguador	0,00690	0,056	5,278	1,170
18	albañil	0,18057	1,278	6,556	26,901
19	alcalde	0,03910	0,306	6,861	6,433
20	alergólogo	0,00299	0,028	6,889	0,585
21	alfarero	0,00224	0,028	6,917	0,585
22	alicatador	0,00314	0,028	6,944	0,585
23	alistador	0,00247	0,028	6,972	0,585
24	almacenista	0,00345	0,028	7,000	0,585
25	ama de casa	0,06896	0,528	7,528	11,111
26	ama de llaves	0,00483	0,056	7,583	1,170
27	analista	0,00845	0,056	7,639	1,170
28	anestesista	0,01272	0,083	7,722	1,754
29	animador	0,00921	0,056	7,778	1,170
30	animador deportivo	0,00259	0,028	7,806	0,585
31	animador social	0,00299	0,028	7,833	0,585
32	antenista	0,00299	0,028	7,861	0,585
33	aparejador	0,03186	0,250	8,111	5,263
34	apicultor	0,01246	0,111	8,222	2,339
35	apoderado	0,00345	0,028	8,250	0,585
36	aprendiz	0,00235	0,028	8,278	0,585
37	aprendiz de mecánico	0,00299	0,028	8,306	0,585
38	árbitro	0,00496	0,056	8,361	1,170
39	armero	0,00204	0,028	8,389	0,585
40	arqueólogo	0,03438	0,250	8,639	5,263
41	arquitecto	0,22886	1,528	10,167	32,164
42	arquitecto técnico	0,01170	0,056	10,222	1,170
43	artesano	0,00235	0,028	10,250	0,585

44	artista	0,07536	0,528	10,778	11,111
45	asesor	0,01925	0,167	10,944	3,509
46	asesor financiero	0,00247	0,028	10,972	0,585
47	asesor fiscal	0,00380	0,028	11,000	0,585
48	asistenta	0,00531	0,028	11,028	0,585
49	asistente	0,02728	0,194	11,222	4,094
50	astronauta	0,01914	0,139	11,361	2,924
51	atleta	0,00932	0,083	11,444	1,754
52	ATS	0,02459	0,222	11,667	4,678
53	au pair	0,00531	0,028	11,694	0,585
54	auditor	0,00439	0,028	11,722	0,585
55	autónomo	0,02143	0,194	11,917	4,094
56	auxiliar	0,01011	0,083	12,000	1,754
57	auxiliar administrativo	0,00520	0,056	12,056	1,170
58	auxiliar de clínica	0,00955	0,083	12,139	1,754
59	auxiliar de enfermería	0,01255	0,111	12,250	2,339
60	aviador	0,00612	0,056	12,306	1,170
61	ayudante de mecánico	0,00314	0,028	12,333	0,585
62	azafato	0,05874	0,444	12,778	9,357
63	azafato de vuelos	0,00194	0,028	12,806	0,585
64	bailador	0,00259	0,028	12,833	0,585
65	bailaor	0,00659	0,056	12,889	1,170
66	bailarín	0,03235	0,333	13,222	7,018
67	banderillero	0,00362	0,028	13,250	0,585
68	banquero	0,08275	0,694	13,944	14,620
69	barman	0,00235	0,028	13,972	0,585
70	barrendero	0,11079	0,833	14,806	17,544
71	barrista	0,00362	0,028	14,833	0,585
72	basurero	0,04951	0,333	15,167	7,018
73	becario	0,01432	0,083	15,250	1,754
74	bedel	0,00345	0,028	15,278	0,585
75	bibliotecario	0,03852	0,278	15,556	5,848
76	biólogo	0,06534	0,583	16,139	12,281
77	bombero	0,19170	1,472	17,611	30,994
78	botones	0,00876	0,083	17,694	1,754
79	boxeador	0,00204	0,028	17,722	0,585
80	bróker	0,02053	0,139	17,861	2,924
81	buceador	0,00235	0,028	17,889	0,585
82	butanero	0,00546	0,056	17,944	1,170
83	buzo	0,00345	0,028	17,972	0,585
84	cabrero	0,01565	0,167	18,139	3,509
85	cajero	0,06039	0,556	18,694	11,696
86	cámara	0,01495	0,111	18,806	2,339
87	camarero	0,24494	1,833	20,639	38,596
88	camarero de pisos	0,00483	0,028	20,667	0,585
89	camionero	0,06509	0,556	21,222	11,696

90	campesino	0,00741	0,083	21,306	1,754
91	canguro	0,01728	0,139	21,444	2,924
92	cantante	0,09050	0,778	22,222	16,374
93	cantaor	0,00851	0,083	22,306	1,754
94	cantautor	0,01377	0,111	22,417	2,339
95	capataz	0,00877	0,056	22,472	1,170
96	carcelero	0,00670	0,056	22,528	1,170
97	cardenal	0,00362	0,056	22,583	1,170
98	cardiólogo	0,00929	0,083	22,667	1,754
99	carnicero	0,11877	0,861	23,528	18,129
100	carpintero	0,13849	0,944	24,472	19,883
101	cartero	0,05207	0,306	24,778	6,433
102	catador	0,00483	0,028	24,806	0,585
103	catedrático	0,00471	0,056	24,861	1,170
104	caudillo	0,00418	0,028	24,889	0,585
105	cazador	0,00285	0,028	24,917	0,585
106	celador	0,05465	0,444	25,361	9,357
107	cerrajero	0,02056	0,167	25,528	3,509
108	chacha	0,00418	0,028	25,556	0,585
109	chapista	0,00971	0,083	25,639	1,754
110	charcutero	0,03563	0,250	25,889	5,263
111	chatarrero	0,00805	0,056	25,944	1,170
112	chef	0,01532	0,111	26,056	2,339
113	chófer	0,04301	0,389	26,444	8,187
114	chófer de camión	0,00247	0,028	26,472	0,585
115	científico	0,02118	0,139	26,611	2,924
116	cineasta	0,00235	0,028	26,639	0,585
117	cirujano	0,02194	0,167	26,806	3,509
118	clarinetista	0,00506	0,028	26,833	0,585
119	cobrador	0,00345	0,028	26,861	0,585
120	cocinero	0,20936	1,639	28,500	34,503
121	comentarista	0,00329	0,028	28,528	0,585
122	comercial	0,03905	0,361	28,889	7,602
123	comerciante	0,01513	0,111	29,000	2,339
124	compositor	0,00857	0,083	29,083	1,754
125	comunicador	0,00362	0,028	29,111	0,585
126	concejal	0,01324	0,111	29,222	2,339
127	concertista	0,00314	0,028	29,250	0,585
128	conductor	0,18248	1,417	30,667	29,825
129	conductor de autobús	0,01210	0,111	30,778	2,339
130	conductor de coches de caballos	0,00380	0,028	30,806	0,585
131	confitero	0,00899	0,056	30,861	1,170
132	consejero	0,00797	0,056	30,917	1,170
133	conserje	0,09552	0,750	31,667	15,789
134	constructor	0,05309	0,361	32,028	7,602
135	consultor	0,00805	0,056	32,083	1,170
136	contable	0,03701	0,278	32,361	5,848

137	contratista	0,00796	0,056	32,417	1,170
138	contrato	0,00259	0,028	32,444	0,585
139	controlador	0,00259	0,028	32,472	0,585
140	controlador aéreo	0,01139	0,111	32,583	2,339
141	coordinador	0,00194	0,028	32,611	0,585
142	coreógrafo	0,00534	0,056	32,667	1,170
143	coronel	0,00604	0,056	32,722	1,170
144	costurero	0,02259	0,194	32,917	4,094
145	creativo	0,00439	0,028	32,944	0,585
146	criado	0,00214	0,028	32,972	0,585
147	cristalero	0,00461	0,056	33,028	1,170
148	cuidador	0,00272	0,028	33,056	0,585
149	cura	0,00931	0,083	33,139	1,754
150	decano	0,00345	0,028	33,167	0,585
151	decorador	0,02739	0,222	33,389	4,678
152	delegado	0,00185	0,028	33,417	0,585
153	delineante	0,01432	0,111	33,528	2,339
154	dentista	0,06668	0,528	34,056	11,111
155	dependiente	0,12836	1,111	35,167	23,392
156	deportista	0,02936	0,222	35,389	4,678
157	deportista de élite	0,00272	0,028	35,417	0,585
158	dermatólogo	0,00399	0,028	35,444	0,585
159	desempleado	0,00345	0,028	35,472	0,585
160	deshollinador	0,00506	0,028	35,500	0,585
161	despido	0,00272	0,028	35,528	0,585
162	detective	0,00506	0,028	35,556	0,585
163	dibujante	0,01295	0,083	35,639	1,754
164	diplomático	0,00483	0,028	35,667	0,585
165	diputado	0,01697	0,111	35,778	2,339
166	director	0,10301	0,889	36,667	18,713
167	director de banco	0,00803	0,056	36,722	1,170
168	director de cine	0,01754	0,139	36,861	2,924
169	director de finanzas	0,00272	0,028	36,889	0,585
170	director de orquesta	0,01571	0,111	37,000	2,339
171	disc-jockey	0,00285	0,028	37,028	0,585
172	diseñador	0,06792	0,556	37,583	11,696
173	diseñador de moda	0,00299	0,028	37,611	0,585
174	diseñador gráfico	0,00439	0,028	37,639	0,585
175	docente	0,03157	0,167	37,806	3,509
176	doctor	0,13465	0,917	38,722	19,298
177	dramaturgo	0,00259	0,028	38,750	0,585
178	ebanista	0,00345	0,028	38,778	0,585
179	ecologista	0,00585	0,028	38,806	0,585
180	economista	0,04436	0,361	39,167	7,602
181	educador	0,00810	0,083	39,250	1,754
182	educador social	0,00878	0,056	39,306	1,170
183	ejecutivo	0,00594	0,056	39,361	1,170
184	electricista	0,17552	1,333	40,694	28,070
185	embajador	0,00235	0,028	40,722	0,585

186	empleado	0,00483	0,028	40,750	0,585
187	empleado de banco	0,01724	0,139	40,889	2,924
188	empleado de hogar	0,00259	0,028	40,917	0,585
189	empresario	0,07924	0,694	41,611	14,620
190	encargado	0,00836	0,056	41,667	1,170
191	encofrador	0,02534	0,194	41,861	4,094
192	endocrino	0,00439	0,028	41,889	0,585
193	enfermero	0,31098	2,222	44,111	46,784
194	enfoscador	0,00439	0,028	44,139	0,585
195	enólogo	0,00483	0,028	44,167	0,585
196	enseñante	0,01487	0,083	44,250	1,754
197	enterrador	0,01441	0,111	44,361	2,339
198	entrenador	0,00805	0,056	44,417	1,170
199	equilibrista	0,00345	0,028	44,444	0,585
200	ertzaintza	0,00362	0,028	44,472	0,585
201	escaparatista	0,00247	0,028	44,500	0,585
202	escenógrafo	0,00299	0,028	44,528	0,585
203	escritor	0,07837	0,583	45,111	12,281
204	escultor	0,03902	0,278	45,389	5,848
205	especulador	0,00668	0,056	45,444	1,170
206	especulador financiero	0,00259	0,028	45,472	0,585
207	espía	0,00557	0,056	45,528	1,170
208	estadista	0,00531	0,028	45,556	0,585
209	estafador	0,00299	0,028	45,583	0,585
210	esteticista	0,01144	0,083	45,667	1,754
211	estilista	0,00658	0,056	45,722	1,170
212	estríper	0,00247	0,028	45,750	0,585
213	estudiante	0,06981	0,556	46,306	11,696
214	eurodiputado	0,00418	0,028	46,333	0,585
215	fabricante	0,00214	0,028	46,361	0,585
216	farmacéutico	0,05934	0,500	46,861	10,526
217	ferrallista	0,00272	0,028	46,889	0,585
218	filólogo	0,04817	0,361	47,250	7,602
219	filósofo	0,01573	0,139	47,389	2,924
220	fiscal	0,02213	0,167	47,556	3,509
221	físico	0,03406	0,278	47,833	5,848
222	fisioterapeuta	0,01852	0,194	48,028	4,094
223	florista	0,01324	0,111	48,139	2,339
224	folclórico	0,00247	0,028	48,167	0,585
225	fontanero	0,19821	1,361	49,528	28,655
226	forense	0,00460	0,028	49,556	0,585
227	fotógrafo	0,03042	0,222	49,778	4,678
228	fraile	0,00224	0,028	49,806	0,585
229	francotirador	0,00439	0,028	49,833	0,585
230	fresador	0,00506	0,028	49,861	0,585
231	friegaplatos	0,00439	0,028	49,889	0,585
232	frigorista	0,00380	0,028	49,917	0,585
233	frutero	0,06628	0,556	50,472	11,696

234	funámbulo	0,00439	0,028	50,500	0,585
235	funcionario	0,04831	0,361	50,861	7,602
236	futbolista	0,07764	0,611	51,472	12,865
237	ganadero	0,03680	0,278	51,750	5,848
238	geógrafo	0,01075	0,083	51,833	1,754
239	geólogo	0,00329	0,028	51,861	0,585
240	gerente	0,02464	0,222	52,083	4,678
241	gestor	0,00959	0,083	52,167	1,754
242	ginecólogo	0,03412	0,278	52,444	5,848
243	gobernador	0,00247	0,028	52,472	0,585
244	gobernanta	0,02159	0,194	52,667	4,094
245	golfista	0,00259	0,028	52,694	0,585
246	granjero	0,01938	0,167	52,861	3,509
247	gruista	0,00798	0,056	52,917	1,170
248	guarda	0,00272	0,028	52,944	0,585
249	guardaespaldas	0,02207	0,194	53,139	4,094
250	guardia	0,01013	0,083	53,222	1,754
251	Guardia Civil	0,04509	0,361	53,583	7,602
252	guardia de seguridad	0,02131	0,167	53,750	3,509
253	guía turístico	0,05137	0,361	54,111	7,602
254	guionista	0,01123	0,083	54,194	1,754
255	guitarrista	0,03111	0,250	54,444	5,263
256	heladero	0,00418	0,028	54,472	0,585
257	herrador	0,00439	0,028	54,500	0,585
258	herrero	0,02575	0,194	54,694	4,094
259	historiador	0,03669	0,278	54,972	5,848
260	hotelero	0,00418	0,028	55,000	0,585
261	huelga	0,00247	0,028	55,028	0,585
262	humorista	0,00272	0,028	55,056	0,585
263	informático	0,11337	0,833	55,889	17,544
264	ingeniero	0,27354	1,778	57,667	37,427
265	ingeniero de caminos	0,00418	0,028	57,694	0,585
266	ingeniero de sistemas	0,00460	0,028	57,722	0,585
267	ingeniero eléctrico	0,00531	0,028	57,750	0,585
268	ingeniero industrial	0,00506	0,028	57,778	0,585
269	ingeniero informático	0,00483	0,028	57,806	0,585
270	ingeniero mecánico	0,00506	0,028	57,833	0,585
271	ingeniero químico	0,00557	0,028	57,861	0,585
272	inmobiliario	0,00153	0,028	57,889	0,585
273	inspector	0,01177	0,083	57,972	1,754
274	interiorista	0,01036	0,083	58,056	1,754
275	intérprete	0,08266	0,583	58,639	12,281
276	interventor	0,00789	0,056	58,694	1,170
277	investigador	0,06455	0,417	59,111	8,772

278	jardinero	0,09557	0,778	59,889	16,374
279	jefe	0,02896	0,222	60,111	4,678
280	jefe de cocina	0,00259	0,028	60,139	0,585
281	jefe de estación	0,00247	0,028	60,167	0,585
282	jefe de mantenimiento	0,00744	0,056	60,222	1,170
283	jefe de obra	0,01285	0,111	60,333	2,339
284	jefe de personal	0,00329	0,028	60,361	0,585
285	jeque árabe	0,00768	0,056	60,417	1,170
286	jornalero	0,00523	0,056	60,472	1,170
287	joyero	0,00765	0,056	60,528	1,170
288	jubilado	0,00418	0,028	60,556	0,585
289	juez	0,11092	0,833	61,389	17,544
290	jugador de baloncesto	0,01516	0,139	61,528	2,924
291	jugador de deporte	0,00299	0,028	61,556	0,585
292	ladrón	0,01648	0,111	61,667	2,339
293	lechero	0,00677	0,056	61,722	1,170
294	librero	0,00914	0,083	61,806	1,754
295	limpiador	0,13249	1,083	62,889	22,807
296	lingüista	0,01336	0,083	62,972	1,754
297	locutor	0,00604	0,056	63,028	1,170
298	logopeda	0,00439	0,028	63,056	0,585
299	maestro	0,26079	1,556	64,611	32,749
300	maestro de escuela	0,00439	0,028	64,639	0,585
301	maestro de obras	0,00235	0,028	64,667	0,585
302	maestro de reiki	0,00214	0,028	64,694	0,585
303	magistrado	0,00701	0,056	64,750	1,170
304	mago	0,00861	0,083	64,833	1,754
305	mâitre	0,00506	0,028	64,861	0,585
306	mánager	0,00628	0,056	64,917	1,170
307	mancebo	0,00399	0,028	64,944	0,585
308	maquetador	0,00285	0,028	64,972	0,585
309	maquillador	0,00399	0,056	65,028	1,170
310	maquinista	0,01579	0,139	65,167	2,924
311	marinero	0,02598	0,194	65,361	4,094
312	marino	0,00408	0,056	65,417	1,170
313	marroquiner	0,00299	0,028	65,444	0,585
314	masajista	0,00804	0,083	65,528	1,754
315	matemático	0,04363	0,361	65,889	7,602
316	matrón	0,00483	0,028	65,917	0,585
317	mecánico	0,06596	0,444	66,361	9,357
318	mecanógrafo	0,00314	0,028	66,389	0,585
319	médico	0,46564	2,972	69,361	62,573
320	médico de cabecera	0,00604	0,056	69,417	1,170
321	médium	0,00345	0,028	69,444	0,585

322	mentor	0,00380	0,028	69,472	0,585
323	mercenario	0,00845	0,056	69,528	1,170
324	militar	0,04301	0,333	69,861	7,018
325	minero	0,00717	0,056	69,917	1,170
326	ministro	0,01793	0,139	70,056	2,924
327	modelo	0,04131	0,333	70,389	7,018
328	modisto	0,00428	0,056	70,444	1,170
329	monaguillo	0,00214	0,028	70,472	0,585
330	monitor	0,01563	0,139	70,611	2,924
331	monja	0,00888	0,083	70,694	1,754
332	monje	0,00204	0,028	70,722	0,585
333	montador	0,00299	0,028	70,750	0,585
334	mozo	0,00194	0,028	70,778	0,585
335	mozo de escuadra	0,00380	0,028	70,806	0,585
336	muletero	0,00380	0,028	70,833	0,585
337	músico	0,11478	0,750	71,583	15,789
338	nadador	0,00259	0,028	71,611	0,585
339	neurólogo	0,00822	0,056	71,667	1,170
340	niñera	0,01744	0,167	71,833	3,509
341	notario	0,03647	0,278	72,111	5,848
342	obispo	0,00389	0,056	72,167	1,170
343	obrero	0,03228	0,222	72,389	4,678
344	oculista	0,00948	0,083	72,472	1,754
345	odontólogo	0,00760	0,056	72,528	1,170
346	oficial	0,00185	0,028	72,556	0,585
347	oficial de la marina	0,00194	0,028	72,583	0,585
348	oficinista	0,02572	0,194	72,778	4,094
349	oftalmólogo	0,00362	0,028	72,806	0,585
350	operador de cámara	0,00399	0,028	72,833	0,585
351	operario	0,00483	0,028	72,861	0,585
352	óptico	0,02135	0,139	73,000	2,924
353	orientador	0,01222	0,083	73,083	1,754
354	ornitólogo	0,00499	0,056	73,139	1,170
355	otorrino	0,00418	0,028	73,167	0,585
356	paisajista	0,00247	0,028	73,194	0,585
357	panadero	0,18823	1,250	74,444	26,316
358	papa	0,00389	0,056	74,500	1,170
359	parado	0,02306	0,139	74,639	2,924
360	paro	0,00557	0,028	74,667	0,585
361	pastelero	0,02609	0,194	74,861	4,094
362	pastor	0,00937	0,083	74,944	1,754
363	patrón	0,00235	0,028	74,972	0,585
364	payaso	0,00967	0,111	75,083	2,339
365	pedagogo	0,02932	0,167	75,250	3,509
366	pediatra	0,02293	0,167	75,417	3,509
367	peletero	0,00439	0,028	75,444	0,585

368	peluquero	0,09481	0,639	76,083	13,450
369	peón	0,04128	0,333	76,417	7,018
370	peón de albañil	0,00644	0,056	76,472	1,170
371	periodista	0,10589	0,833	77,306	17,544
372	perito	0,01647	0,139	77,444	2,924
373	personal doméstico	0,00224	0,028	77,472	0,585
374	pescadero	0,08091	0,667	78,139	14,035
375	pescador	0,02918	0,250	78,389	5,263
376	pianista	0,02148	0,167	78,556	3,509
377	piloto	0,07158	0,611	79,167	12,865
378	pinche	0,01357	0,111	79,278	2,339
379	pintor	0,17975	1,222	80,500	25,731
380	pirata informático	0,00285	0,028	80,528	0,585
381	podólogo	0,00876	0,083	80,611	1,754
382	poeta	0,01416	0,111	80,722	2,339
383	policía	0,29559	2,167	82,889	45,614
384	político	0,11035	0,889	83,778	18,713
385	porquero	0,00247	0,028	83,806	0,585
386	portero	0,00994	0,111	83,917	2,339
387	presentador	0,05614	0,500	84,417	10,526
388	presidente	0,01699	0,167	84,583	3,509
389	probador de productos	0,00299	0,028	84,611	0,585
390	procurador	0,01671	0,111	84,722	2,339
391	productor	0,01481	0,111	84,833	2,339
392	profesor	0,78858	4,306	89,139	90,643
393	profesor de secundaria	0,00460	0,028	89,167	0,585
394	profesor de yoga	0,00224	0,028	89,194	0,585
395	profesor universitario	0,00418	0,028	89,222	0,585
396	programador	0,02213	0,167	89,389	3,509
397	promotor	0,00259	0,028	89,417	0,585
398	prostituto	0,01822	0,139	89,556	2,924
399	proxeneta	0,00224	0,028	89,583	0,585
400	psicoanalista	0,00235	0,028	89,611	0,585
401	psicólogo	0,10662	0,806	90,417	16,959
402	psiquiatra	0,03858	0,306	90,722	6,433
403	publicista	0,02251	0,139	90,861	2,924
404	químico	0,06470	0,472	91,333	9,942
405	quiosquero	0,01244	0,111	91,444	2,339
406	quiromasajista	0,00259	0,028	91,472	0,585
407	realizador	0,00460	0,028	91,500	0,585
408	recadero	0,00329	0,028	91,528	0,585
409	repcionista	0,05737	0,472	92,000	9,942
410	recolector de frutos	0,00380	0,028	92,028	0,585
411	rector	0,00146	0,028	92,056	0,585

412	redactor	0,00362	0,028	92,083	0,585
413	relaciones públicas	0,00314	0,028	92,111	0,585
414	repartidor	0,01881	0,167	92,278	3,509
415	reponedor	0,01761	0,194	92,472	4,094
416	reportero	0,00185	0,028	92,500	0,585
417	restaurador	0,00742	0,056	92,556	1,170
418	revisor	0,00532	0,056	92,611	1,170
419	sargento	0,00235	0,028	92,639	0,585
420	sastre	0,02937	0,222	92,861	4,678
421	secretario	0,08128	0,667	93,528	14,035
422	sereno	0,00399	0,028	93,556	0,585
423	sicario	0,00771	0,056	93,611	1,170
424	sirviente	0,00658	0,056	93,667	1,170
425	socorrista	0,00952	0,083	93,750	1,754
426	solador	0,00299	0,028	93,778	0,585
427	soldado	0,01990	0,139	93,917	2,924
428	soldador	0,01743	0,139	94,056	2,924
429	solista	0,00299	0,028	94,083	0,585
430	subdirector	0,00345	0,028	94,111	0,585
431	sumiller	0,01177	0,083	94,194	1,754
432	taquillero	0,00259	0,028	94,222	0,585
433	tatuador	0,00314	0,028	94,250	0,585
434	taxidermista	0,01214	0,111	94,361	2,339
435	taxista	0,15530	1,167	95,528	24,561
436	técnico	0,01002	0,083	95,611	1,754
437	técnico de imagen y sonido	0,00224	0,028	95,639	0,585
438	técnico de rayos	0,00259	0,028	95,667	0,585
439	tejedor	0,00418	0,028	95,694	0,585
440	telefonista	0,00927	0,083	95,778	1,754
441	teleoperador	0,02306	0,194	95,972	4,094
442	tendero	0,04939	0,361	96,333	7,602
443	teniente alcalde	0,00259	0,028	96,361	0,585
444	tenista	0,00792	0,083	96,444	1,754
445	tenor	0,00285	0,028	96,472	0,585
446	terapeuta	0,00299	0,028	96,500	0,585
447	terminólogo	0,00506	0,028	96,528	0,585
448	timador	0,00627	0,056	96,583	1,170
449	top-model	0,00914	0,083	96,667	1,754
450	topógrafo	0,01667	0,111	96,778	2,339
451	torero	0,03613	0,278	97,056	5,848
452	tornero	0,00531	0,028	97,083	0,585
453	trabajo	0,00585	0,028	97,111	0,585
454	traductor	0,13550	0,917	98,028	19,298
455	traductor jurado	0,00483	0,028	98,056	0,585
456	transportista	0,01125	0,111	98,167	2,339

457	traumatólogo	0,01006	0,083	98,250	1,754
458	trompetista	0,00840	0,056	98,306	1,170
459	trovador	0,00247	0,028	98,333	0,585
460	tutor	0,01115	0,056	98,389	1,170
461	vaquero	0,00235	0,028	98,417	0,585
462	vendedor	0,07132	0,556	98,972	11,696
463	verdugo	0,00460	0,028	99,000	0,585
464	verdulero	0,00743	0,083	99,083	1,754
465	veterinario	0,01594	0,111	99,194	2,339
466	vidente	0,00362	0,028	99,222	0,585
467	vigilante	0,01246	0,111	99,333	2,339
468	vigilante de la playa	0,00314	0,028	99,361	0,585
469	vinicultor	0,00439	0,028	99,389	0,585
470	violinista	0,02189	0,139	99,528	2,924
471	violonchelista	0,00557	0,028	99,556	0,585
472	visitador médico	0,00438	0,056	99,611	1,170
473	vulcanólogo	0,00362	0,028	99,639	0,585
474	yesero	0,00460	0,028	99,667	0,585
475	yesista	0,00259	0,028	99,694	0,585
476	zapador	0,00329	0,028	99,722	0,585
477	zapatero	0,02940	0,250	99,972	5,263
478	zoólogo	0,00177	0,028	100,000	0,585

1.1.9. Ordenadores e internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(Microsoft) Access	0,00708	0,061	0,061	1,163
2	(Microsoft) Excel	0,02436	0,212	0,273	4,070
3	(Microsoft) Internet Explorer	0,01763	0,152	0,425	2,907
4	(Microsoft) Office	0,00725	0,061	0,486	1,163
5	(Microsoft) PowerPoint	0,03120	0,273	0,759	5,233
6	(Microsoft) Word	0,05469	0,455	1,214	8,721
7	acceder	0,00145	0,030	1,244	0,581
8	acceso	0,00223	0,030	1,275	0,581
9	accesorio	0,00297	0,030	1,305	0,581
10	aceptar	0,00245	0,030	1,335	0,581
11	Acer	0,00528	0,030	1,366	0,581
12	actualizar	0,00430	0,061	1,426	1,163
13	adaptador	0,00270	0,030	1,457	0,581
14	adaptador inalámbrico	0,00343	0,030	1,487	0,581
15	adjuntar	0,00813	0,091	1,578	1,744
16	administrador	0,00184	0,030	1,608	0,581
17	administrador de tareas	0,00480	0,030	1,639	0,581
18	Adobe Photoshop	0,00519	0,061	1,700	1,163
19	ADSL	0,03723	0,303	2,003	5,814
20	agregar contacto	0,00193	0,030	2,033	0,581
21	Ajax	0,00554	0,030	2,064	0,581
22	alfombrilla	0,04138	0,303	2,367	5,814
23	almacenamiento	0,00583	0,061	2,428	1,163
24	altavoz	0,09101	0,698	3,126	13,372
25	amistad	0,00184	0,030	3,156	0,581
26	amplitud	0,00223	0,030	3,187	0,581
27	ancho de banda	0,00396	0,030	3,217	0,581
28	Android	0,00343	0,030	3,247	0,581
29	antena	0,00343	0,030	3,278	0,581
30	antiespía	0,00203	0,030	3,308	0,581
31	antivirus	0,07319	0,698	4,006	13,372
32	aplicación	0,01449	0,152	4,158	2,907
33	Apple	0,01995	0,152	4,310	2,907
34	árbol	0,00312	0,030	4,340	0,581
35	archivo	0,16377	1,305	5,645	25,000
36	Ares	0,00437	0,061	5,706	1,163
37	arreglo	0,00159	0,030	5,736	0,581
38	Audacity	0,00360	0,030	5,766	0,581
39	audio	0,00327	0,030	5,797	0,581
40	aula virtual	0,00581	0,030	78,270	0,581
41	auricular	0,00396	0,030	5,827	0,581

42	AutoCAD	0,01388	0,121	5,948	2,326
43	Avast	0,00312	0,030	5,979	0,581
44	ayuda	0,00245	0,030	6,009	0,581
45	bajar	0,01463	0,091	6,100	1,744
46	bajar archivo	0,01542	0,121	6,222	2,326
47	bajar música	0,00416	0,030	6,252	0,581
48	banda ancha	0,01834	0,121	6,373	2,326
49	bandeja de entrada	0,00915	0,121	6,495	2,326
50	barra	0,00360	0,030	6,525	0,581
51	barra de herramientas	0,00938	0,091	6,616	1,744
52	base de datos	0,02190	0,182	6,798	3,488
53	batería	0,05597	0,455	7,253	8,721
54	BIN	0,00378	0,030	7,284	0,581
55	bit	0,00234	0,030	7,314	0,581
56	BitTorrent	0,00446	0,061	7,375	1,163
57	blog	0,11262	0,880	8,255	16,860
58	bluetooth	0,00564	0,061	8,316	1,163
59	borrar	0,00343	0,030	8,346	0,581
60	botón	0,00674	0,091	8,437	1,744
61	botón derecho	0,00257	0,030	8,467	0,581
62	burbuja	0,00184	0,030	8,498	0,581
63	buscador	0,13783	1,153	9,651	22,093
64	buscar	0,01097	0,091	9,742	1,744
65	buscar en la red	0,00378	0,030	9,772	0,581
66	buscar información	0,00831	0,061	9,833	1,163
67	búsqueda	0,00378	0,030	9,863	0,581
68	byte	0,01023	0,091	9,954	1,744
69	C++	0,00504	0,030	9,985	0,581
70	cable	0,17350	1,366	11,351	26,163
71	cable de red	0,00312	0,030	11,381	0,581
72	cable de serie	0,00396	0,030	11,411	0,581
73	cable USB	0,00312	0,030	11,442	0,581
74	caja	0,00624	0,061	11,502	1,163
75	calendario	0,00343	0,030	11,533	0,581
76	cámara	0,01017	0,091	11,624	1,744
77	cámara web	0,03160	0,303	11,927	5,814
78	campus	0,00458	0,030	11,958	0,581
79	campus virtual	0,04384	0,303	12,261	5,814
80	cargador	0,03589	0,273	12,534	5,233
81	cargar	0,01085	0,061	12,595	1,163
82	carpeta	0,08567	0,728	13,323	13,953
83	cartucho de tóner	0,00396	0,030	13,354	0,581
84	casco	0,01155	0,091	13,445	1,744
85	casete	0,00416	0,030	13,475	0,581
86	casilla	0,00283	0,030	13,505	0,581
87	CD(-ROM)	0,09981	0,850	14,355	16,279
88	centro TIC	0,00312	0,030	14,385	0,581
89	certificado	0,00120	0,030	14,416	0,581

90	chat	0,12461	1,062	15,478	20,349
91	chatear	0,07561	0,577	16,055	11,047
92	chip	0,00693	0,061	16,115	1,163
93	cibernauta	0,00184	0,030	16,146	0,581
94	cine	0,00312	0,030	16,176	0,581
95	clave	0,01014	0,091	16,267	1,744
96	clavija	0,00213	0,030	16,297	0,581
97	clic	0,01407	0,121	16,419	2,326
98	cliente	0,00527	0,061	16,480	1,163
99	codificación	0,00145	0,030	16,510	0,581
100	código	0,00581	0,030	16,540	0,581
101	color	0,00514	0,061	16,601	1,163
102	comando	0,00473	0,061	16,662	1,163
103	compilador	0,01475	0,121	16,783	2,326
104	compilar	0,00562	0,061	16,844	1,163
105	comprimir	0,00396	0,030	16,874	0,581
106	computadora	0,01586	0,091	16,965	1,744
107	comunicación	0,01808	0,152	17,117	2,907
108	condensador	0,00541	0,061	17,178	1,163
109	conectarse	0,00436	0,030	17,208	0,581
110	conectividad	0,00480	0,030	17,238	0,581
111	conector	0,00982	0,091	17,329	1,744
112	conector DIN	0,00257	0,030	17,360	0,581
113	conexión	0,09108	0,789	18,149	15,116
114	conexión ADSL	0,00245	0,030	18,179	0,581
115	conexión telefónica	0,00416	0,030	18,209	0,581
116	configuración	0,00416	0,030	18,240	0,581
117	configurar	0,00749	0,091	18,331	1,744
118	conmutador	0,00360	0,030	18,361	0,581
119	conocer gente	0,00270	0,030	18,392	0,581
120	consola	0,00327	0,030	18,422	0,581
121	consulta	0,00234	0,030	18,452	0,581
122	contacto	0,00396	0,030	18,483	0,581
123	contenido	0,00297	0,030	18,513	0,581
124	contraseña	0,03636	0,273	18,786	5,233
125	control parental	0,00343	0,030	18,816	0,581
126	control remoto	0,00203	0,030	18,847	0,581
127	controlador	0,00840	0,061	18,907	1,163
128	cookie	0,02416	0,243	19,150	4,651
129	copiadora	0,00327	0,030	19,181	0,581
130	copiar	0,01427	0,152	19,332	2,907
131	copiar documento	0,00396	0,030	19,363	0,581
132	CorelDRAW	0,00769	0,061	19,423	1,163
133	correo (electrónico)	0,24676	1,973	21,396	37,791
134	corriente	0,00528	0,030	21,426	0,581
135	cortafuego	0,05172	0,455	21,882	8,721
136	cortar	0,00656	0,061	21,942	1,163
137	CPU	0,04878	0,364	22,307	6,977
138	cristal de cuarzo	0,00312	0,030	22,337	0,581

139	cultura	0,00193	0,030	22,367	0,581
140	cursor	0,00480	0,030	22,398	0,581
141	dato	0,00416	0,030	22,428	0,581
142	Dell	0,00852	0,061	22,489	1,163
143	depurar	0,00175	0,030	22,519	0,581
144	descarga	0,06511	0,577	23,096	11,047
145	descarga directa	0,00159	0,030	23,126	0,581
146	descargar	0,07546	0,607	23,733	11,628
147	descargar archivo	0,02755	0,182	23,915	3,488
148	descargar documento	0,00504	0,030	23,945	0,581
149	descompilar	0,00343	0,030	23,976	0,581
150	descomprimir	0,00378	0,030	24,006	0,581
151	desfragmentar	0,01086	0,091	24,097	1,744
152	desinstalar	0,00625	0,061	24,158	1,163
153	diapositiva	0,00283	0,030	24,188	0,581
154	diccionario	0,00270	0,030	24,219	0,581
155	dígito	0,00312	0,030	24,249	0,581
156	dirección	0,01010	0,121	24,370	2,326
157	dirección de correo (electrónico)	0,00762	0,061	24,431	1,163
158	dirección web	0,00982	0,091	24,522	1,744
159	directorio	0,01160	0,121	24,643	2,326
160	disco	0,03455	0,273	24,917	5,233
161	disco compacto	0,00378	0,030	24,947	0,581
162	disco duro	0,24014	1,942	26,889	37,209
163	disco duro externo	0,01146	0,121	27,011	2,326
164	disco duro externo multimedia	0,00257	0,030	27,041	0,581
165	disco extraíble	0,00343	0,030	27,071	0,581
166	disco portátil	0,00297	0,030	27,102	0,581
167	disco regrabable	0,00416	0,030	27,132	0,581
168	diseño gráfico	0,00464	0,061	27,193	1,163
169	dispositivo	0,00223	0,030	27,223	0,581
170	disquete	0,03389	0,273	27,496	5,233
171	disquetera	0,00535	0,061	27,557	1,163
172	distracción	0,00360	0,030	27,587	0,581
173	DNS	0,00524	0,061	27,648	1,163
174	doble clic	0,00145	0,030	27,678	0,581
175	documental	0,00245	0,030	27,709	0,581
176	documento	0,05840	0,516	28,225	9,884
177	documento de texto	0,00360	0,030	28,255	0,581
178	dominio	0,00297	0,030	28,285	0,581
179	dominio web	0,00312	0,030	28,316	0,581
180	Dropbox	0,00360	0,030	28,346	0,581
181	DVD	0,06405	0,546	28,892	10,465
182	editor de texto	0,00297	0,030	28,923	0,581
183	educación	0,00297	0,030	28,953	0,581
184	ejecutable	0,00343	0,030	28,983	0,581

185	ejecutar	0,00956	0,091	29,074	1,744
186	El Tiempo	0,00562	0,061	29,135	1,163
187	electricidad	0,00458	0,030	29,165	0,581
188	electrónica	0,00283	0,030	29,196	0,581
189	e-mail	0,05202	0,425	29,621	8,140
190	eMule	0,01665	0,182	29,803	3,488
191	enchufe	0,03287	0,273	30,076	5,233
192	enciclopedia	0,00257	0,030	30,106	0,581
193	enciclopedia virtual	0,00458	0,030	30,137	0,581
194	enlace	0,03353	0,303	30,440	5,814
195	ensamblador	0,00245	0,030	30,470	0,581
196	entretenimiento	0,00436	0,030	30,501	0,581
197	enviar	0,00203	0,030	30,531	0,581
198	error	0,00184	0,030	30,561	0,581
199	escanear	0,00832	0,061	30,622	1,163
200	escáner	0,13177	1,093	31,715	20,930
201	escribir	0,00360	0,030	31,745	0,581
202	escritorio	0,04821	0,425	32,170	8,140
203	espacio virtual	0,00312	0,030	32,200	0,581
204	estilo	0,00257	0,030	32,231	0,581
205	Ethernet	0,00245	0,030	32,261	0,581
206	Europa FM	0,00554	0,030	32,291	0,581
207	extensión	0,00327	0,030	32,322	0,581
208	extraíble	0,00436	0,030	32,352	0,581
209	Facebook	0,20834	1,608	33,961	30,814
210	Farm Ville	0,00480	0,030	33,991	0,581
211	favorito	0,00396	0,061	34,052	1,163
212	fax	0,00480	0,030	34,082	0,581
213	fichero	0,01736	0,121	34,203	2,326
214	firewall	0,00213	0,030	34,234	0,581
215	FireWire	0,00270	0,030	34,264	0,581
216	fondo	0,00234	0,030	34,294	0,581
217	fondo de pantalla	0,00360	0,030	34,325	0,581
218	formatear	0,04203	0,334	34,659	6,395
219	formato	0,00245	0,030	34,689	0,581
220	formato de negrita	0,00270	0,030	34,719	0,581
221	formato jpg	0,00491	0,061	34,780	1,163
222	foro	0,06165	0,516	35,296	9,884
223	freaky	0,00581	0,030	35,326	0,581
224	FriendFeed	0,00458	0,030	35,357	0,581
225	FTP	0,00270	0,030	35,387	0,581
226	fuelle de alimentación	0,01231	0,121	35,508	2,326
227	función	0,00480	0,030	35,539	0,581
228	Gateway	0,00343	0,030	35,569	0,581
229	gigabyte	0,00758	0,091	35,660	1,744
230	global	0,00436	0,030	35,690	0,581
231	Gmail	0,01360	0,121	35,812	2,326
232	GNU	0,00458	0,030	35,842	0,581

233	Google	0,14110	1,184	37,026	22,674
234	Google Crome	0,00283	0,030	37,056	0,581
235	Google Earth	0,00436	0,030	37,086	0,581
236	Google Maps	0,00639	0,061	37,147	1,163
237	Google Reader	0,00436	0,030	37,178	0,581
238	GPRS	0,00312	0,030	37,208	0,581
239	grabador de CD	0,00234	0,030	37,238	0,581
240	grabadora	0,04770	0,395	37,633	7,558
241	grabar	0,00960	0,091	37,724	1,744
242	Gran Hermano	0,00528	0,030	37,754	0,581
243	grupo internauta	0,00436	0,030	37,785	0,581
244	Guadalinx	0,00343	0,030	37,815	0,581
245	guardar	0,01440	0,121	37,936	2,326
246	guardar documento	0,00378	0,030	37,967	0,581
247	gusano	0,00343	0,030	37,997	0,581
248	hacer clic	0,00270	0,030	38,027	0,581
249	hackear	0,00283	0,030	38,058	0,581
250	hacker	0,00378	0,030	38,088	0,581
251	hardware	0,15544	1,244	39,332	23,837
252	HDMI	0,00132	0,030	39,363	0,581
253	herramienta	0,00612	0,061	39,423	1,163
254	hibernar	0,00257	0,030	39,454	0,581
255	hipervínculo	0,00297	0,030	39,484	0,581
256	historial	0,00867	0,091	39,575	1,744
257	Hofmann	0,00283	0,030	39,605	0,581
258	hoja de cálculo	0,00654	0,061	39,666	1,163
259	Hot Potatoes	0,03059	0,273	39,939	5,233
260	Hotmail	0,05771	0,546	40,486	10,465
261	HP	0,00416	0,030	40,516	0,581
262	HTML	0,00480	0,030	40,546	0,581
263	icono	0,01344	0,152	40,698	2,907
264	imagen	0,00642	0,061	40,759	1,163
265	impresora	0,20628	1,548	42,307	29,651
266	imprimir	0,01178	0,091	42,398	1,744
267	inalámbrico	0,02758	0,273	42,671	5,233
268	InfoJobs	0,00458	0,030	42,701	0,581
269	información	0,03225	0,273	42,974	5,233
270	informática	0,00519	0,061	43,035	1,163
271	informático	0,00436	0,030	43,065	0,581
272	ingeniero	0,00297	0,030	43,096	0,581
273	inicio	0,00396	0,030	43,126	0,581
274	inscripción	0,00297	0,030	43,156	0,581
275	insertar	0,00472	0,061	43,217	1,163
276	instalación	0,00334	0,061	43,278	1,163
277	instalar	0,02645	0,243	43,520	4,651
278	instrucción	0,00451	0,061	43,581	1,163
279	Intel Core	0,00528	0,030	43,612	0,581

280	interactivo	0,00234	0,030	43,642	0,581
281	intercambio	0,00175	0,030	43,672	0,581
282	interfaz	0,00550	0,061	43,733	1,163
283	interminable	0,00213	0,030	43,763	0,581
284	internet	0,08004	0,577	44,340	11,047
285	Interplex	0,00257	0,030	44,370	0,581
286	intérprete de comandos	0,00480	0,030	44,401	0,581
287	interrupción	0,00203	0,030	44,431	0,581
288	intranet	0,00944	0,091	44,522	1,744
289	Intro	0,00297	0,030	44,552	0,581
290	inventor	0,00193	0,030	44,583	0,581
291	IP	0,01396	0,121	44,704	2,326
292	iPad	0,00922	0,091	44,795	1,744
293	iPhone	0,00360	0,030	44,825	0,581
294	iPod	0,00213	0,030	44,856	0,581
295	Java	0,01322	0,091	44,947	1,744
296	JClic	0,01993	0,182	45,129	3,488
297	juego	0,03272	0,303	45,432	5,814
298	juego online	0,00815	0,061	45,493	1,163
299	lápiz de memoria	0,01572	0,152	45,645	2,907
300	lápiz óptico	0,00855	0,061	45,706	1,163
301	lápiz táctil	0,00234	0,030	45,736	0,581
302	lápiz USB	0,00504	0,030	45,766	0,581
303	lector de CD-ROM	0,00283	0,030	45,797	0,581
304	lengua	0,00184	0,030	45,827	0,581
305	lenguaje de programación	0,00528	0,030	45,857	0,581
306	lenguaje informático	0,00840	0,061	45,918	1,163
307	letra	0,00458	0,030	45,948	0,581
308	LibreOffice	0,00436	0,030	45,979	0,581
309	librería	0,00458	0,030	46,009	0,581
310	libro electrónico	0,00378	0,030	46,039	0,581
311	licencia	0,00312	0,030	46,070	0,581
312	línea	0,00297	0,030	46,100	0,581
313	LinkedIn	0,00480	0,030	46,131	0,581
314	Linux	0,03095	0,243	46,373	4,651
315	Macintosh	0,02218	0,182	46,555	3,488
316	mando a distancia	0,00416	0,030	46,586	0,581
317	mantenimiento	0,00312	0,030	46,616	0,581
318	manual	0,00416	0,030	46,646	0,581
319	marcador	0,00503	0,061	46,707	1,163
320	máscara	0,00213	0,030	46,737	0,581
321	mecanografía	0,00360	0,030	46,768	0,581
322	medio de comunicación social	0,00312	0,030	46,798	0,581
323	mega	0,01303	0,121	46,920	2,326

324	megabyte	0,01199	0,121	47,041	2,326
325	megapíxel	0,00430	0,061	47,102	1,163
326	memoria	0,08892	0,698	47,800	13,372
327	memoria externa	0,00807	0,061	47,860	1,163
328	memoria interna	0,00873	0,061	47,921	1,163
329	memoria RAM	0,05349	0,516	48,437	9,884
330	memoria USB	0,00378	0,030	48,467	0,581
331	mensaje	0,01349	0,121	48,589	2,326
332	mensajería	0,00605	0,061	48,649	1,163
333	mesa	0,00327	0,030	48,680	0,581
334	Messenger	0,04280	0,364	49,044	6,977
335	Mi PC	0,00351	0,061	49,105	1,163
336	micrófono	0,00480	0,030	49,135	0,581
337	microprocesador	0,01739	0,182	49,317	3,488
338	Microsoft	0,02619	0,212	49,530	4,070
339	Mis documentos	0,00360	0,030	49,560	0,581
340	módem	0,08891	0,789	50,349	15,116
341	monitor	0,19647	1,335	51,684	25,581
342	Moodle	0,01175	0,091	51,775	1,744
343	Mozilla Firefox	0,00514	0,061	51,836	1,163
344	MP3	0,00343	0,030	51,866	0,581
345	MS-DOS	0,00581	0,030	51,897	0,581
346	MSN	0,01243	0,121	52,018	2,326
347	multifunción	0,00343	0,030	52,049	0,581
348	multimedia	0,00203	0,030	52,079	0,581
349	música	0,02939	0,243	52,322	4,651
350	navegador	0,04230	0,364	52,686	6,977
351	navegar	0,07005	0,546	53,232	10,465
352	netbook	0,02446	0,152	53,384	2,907
353	nombre de usuario	0,00554	0,030	53,414	0,581
354	notebook	0,01557	0,121	53,536	2,326
355	noticia	0,00748	0,091	53,627	1,744
356	nuevas tecnologías	0,00327	0,030	53,657	0,581
357	objeto	0,00436	0,030	53,687	0,581
358	ofimática	0,00487	0,061	53,748	1,163
359	online	0,00245	0,030	53,778	0,581
360	operador de búsqueda	0,00297	0,030	53,809	0,581
361	Orange	0,00213	0,030	53,839	0,581
362	ordenador	0,09654	0,546	54,385	10,465
363	ordenador de sobremesa	0,03314	0,212	54,598	4,070
364	ordenador personal	0,01163	0,061	54,659	1,163
365	output	0,00257	0,030	54,689	0,581
366	página	0,02362	0,182	54,871	3,488
367	página de inicio	0,00312	0,030	54,901	0,581

368	página oficial	0,00480	0,030	54,932	0,581
369	página principal	0,00613	0,061	54,992	1,163
370	página web	0,28249	2,124	57,117	40,698
371	página web de contenido erótico	0,00458	0,030	57,147	0,581
372	Paint	0,00803	0,091	57,238	1,744
373	palabra	0,00223	0,030	57,269	0,581
374	palanca de mando	0,00609	0,061	57,329	1,163
375	panel de control	0,00270	0,030	57,360	0,581
376	pantalla	0,61087	3,733	61,093	71,512
377	papelera	0,00396	0,030	61,123	0,581
378	papelera de reciclaje	0,00601	0,061	61,184	1,163
379	paralelismo	0,00175	0,030	61,214	0,581
380	parque tecnológico	0,00283	0,030	61,244	0,581
381	PC	0,04703	0,273	61,517	5,233
382	PDA	0,00193	0,030	61,548	0,581
383	pdf	0,01432	0,152	61,700	2,907
384	pegar	0,01107	0,121	61,821	2,326
385	película	0,02236	0,212	62,033	4,070
386	pendrive	0,15910	1,275	63,308	24,419
387	Pentium	0,00416	0,030	63,338	0,581
388	perfil	0,00203	0,030	63,369	0,581
389	periférico	0,00999	0,091	63,460	1,744
390	periódico	0,00458	0,030	63,490	0,581
391	periódico virtual	0,00312	0,030	63,520	0,581
392	pestaña	0,00360	0,030	63,551	0,581
393	Picassa	0,00297	0,030	63,581	0,581
394	pinchar	0,00609	0,061	63,642	1,163
395	pirata informático	0,00756	0,061	63,703	1,163
396	piratear	0,01358	0,121	63,824	2,326
397	píxel	0,00327	0,030	63,854	0,581
398	placa base	0,07313	0,577	64,431	11,047
399	placa madre	0,00755	0,061	64,492	1,163
400	podcast	0,00283	0,030	64,522	0,581
401	portafolio	0,00257	0,030	64,552	0,581
402	portal	0,01808	0,152	64,704	2,907
403	portal web	0,00203	0,030	64,734	0,581
404	portátil	0,26578	1,760	66,495	33,721
405	prensa	0,00145	0,030	66,525	0,581
406	problema	0,00167	0,030	66,555	0,581
407	procesador	0,02463	0,243	66,798	4,651
408	procesador de textos	0,05344	0,455	67,253	8,721
409	proceso	0,00573	0,061	67,314	1,163
410	programa	0,20668	1,700	69,014	32,558
411	programa informático	0,01232	0,091	69,105	1,744

412	programa operativo	0,00378	0,030	69,135	0,581
413	programación televisiva	0,00312	0,030	69,165	0,581
414	programador	0,00853	0,091	69,256	1,744
415	programar	0,01288	0,152	69,408	2,907
416	protocolo	0,01097	0,121	69,530	2,326
417	Protocolo http	0,00360	0,030	69,560	0,581
418	proveedor	0,00378	0,030	69,590	0,581
419	Proxy	0,00416	0,030	69,621	0,581
420	proyector	0,00814	0,061	69,681	1,163
421	publicidad	0,00203	0,030	69,712	0,581
422	puerto	0,00446	0,061	69,772	1,163
423	puerto USB	0,05000	0,425	70,197	8,140
424	puesto	0,00270	0,030	70,228	0,581
425	puntero	0,00546	0,061	70,288	1,163
426	raíz	0,00458	0,030	70,319	0,581
427	RAM	0,01008	0,061	70,379	1,163
428	ranura	0,00193	0,030	70,410	0,581
429	rapidez	0,00167	0,030	70,440	0,581
430	ratón	0,68471	4,067	74,507	77,907
431	ratón inalámbrico	0,00654	0,061	74,568	1,163
432	red	0,08228	0,637	75,205	12,209
433	red 2.0	0,00360	0,030	75,235	0,581
434	red de alta velocidad	0,00378	0,030	75,266	0,581
435	red inalámbrica	0,02386	0,212	75,478	4,070
436	red social	0,21344	1,730	77,208	33,140
437	reportaje	0,00360	0,030	77,238	0,581
438	resistencia	0,00327	0,030	77,269	0,581
439	resolución	0,00490	0,061	77,329	1,163
440	retardo	0,00283	0,030	77,360	0,581
441	Rincón del Vago	0,00270	0,030	77,390	0,581
442	RIP	0,00396	0,030	77,420	0,581
443	router	0,08640	0,759	78,179	14,535
444	RSS	0,00270	0,030	78,209	0,581
445	rutina	0,00396	0,030	78,240	0,581
446	Safari	0,00600	0,061	78,331	1,163
447	salvapantalla	0,01992	0,152	78,483	2,907
448	script	0,00270	0,030	78,513	0,581
449	SDL Trados	0,00245	0,030	78,543	0,581
450	seguridad	0,00351	0,061	78,604	1,163
451	seleccionar	0,00283	0,030	78,634	0,581
452	serie	0,00297	0,030	78,665	0,581
453	serie online	0,00167	0,030	78,695	0,581
454	Series Yonkis	0,01024	0,091	78,786	1,744
455	servidor	0,06827	0,607	79,393	11,628
456	servidor web	0,00223	0,030	79,423	0,581

457	silla	0,00312	0,030	79,454	0,581
458	simulador de circuitos	0,00436	0,030	79,484	0,581
459	Sin la Mula	0,00234	0,030	79,514	0,581
460	sistema	0,00468	0,061	79,575	1,163
461	sistema binario	0,00297	0,030	79,605	0,581
462	sistema operativo	0,04577	0,364	79,970	6,977
463	sitio	0,00327	0,030	80,000	0,581
464	sitio web	0,02732	0,212	80,212	4,070
465	Skype	0,01154	0,091	80,303	1,744
466	SlideShare	0,00396	0,030	80,334	0,581
467	social	0,00203	0,030	80,364	0,581
468	software	0,16137	1,305	81,669	25,000
469	software libre	0,00312	0,030	81,700	0,581
470	solitario	0,00327	0,030	81,730	0,581
471	sonido	0,00436	0,030	81,760	0,581
472	Sony Vaio	0,00458	0,030	81,791	0,581
473	spam	0,00396	0,061	81,851	1,163
474	Spotify	0,01136	0,091	81,942	1,744
475	streaming	0,00708	0,061	82,003	1,163
476	subir	0,02726	0,243	82,246	4,651
477	subir archivos	0,00927	0,091	82,337	1,744
478	surfear	0,01533	0,121	82,458	2,326
479	suscripción	0,00297	0,030	82,489	0,581
480	tableta	0,00257	0,030	82,519	0,581
481	tableta digitalizadora	0,00721	0,061	82,580	1,163
482	tableta gráfica	0,00360	0,030	82,610	0,581
483	tarifa plana	0,00245	0,030	82,640	0,581
484	tarjeta	0,00496	0,061	82,701	1,163
485	tarjeta de audio	0,00480	0,030	82,731	0,581
486	tarjeta de memoria	0,01505	0,121	82,853	2,326
487	tarjeta de red	0,00785	0,061	82,914	1,163
488	tarjeta de sonido	0,01233	0,091	83,005	1,744
489	tarjeta de vídeo	0,00670	0,061	83,065	1,163
490	tarjeta gráfica	0,05560	0,486	83,551	9,302
491	tarjeta sintonizadora de televisión	0,00396	0,030	83,581	0,581
492	TCP	0,00203	0,030	83,612	0,581
493	tecla	0,06655	0,486	84,097	9,302
494	teclado	0,66352	4,097	88,194	78,488
495	teclado numérico	0,00396	0,030	88,225	0,581
496	técnico	0,00816	0,091	88,316	1,744
497	Telefónica	0,00213	0,030	88,346	0,581
498	teléfono	0,00490	0,061	88,407	1,163
499	teleformación	0,00234	0,030	88,437	0,581
500	teletrabajo	0,00223	0,030	88,467	0,581

501	televisión	0,00223	0,030	88,498	0,581
502	tensión	0,00480	0,030	88,528	0,581
503	terabyte	0,00159	0,030	88,558	0,581
504	terminal virtual	0,00416	0,030	88,589	0,581
505	Terra	0,00234	0,030	88,619	0,581
506	TIC	0,02348	0,152	88,771	2,907
507	tinta	0,00778	0,091	88,862	1,744
508	tóner	0,00416	0,030	88,892	0,581
509	torre	0,22941	1,700	90,592	32,558
510	Toshiba	0,01347	0,091	90,683	1,744
511	trabajo	0,00430	0,061	90,744	1,163
512	transformador	0,00436	0,030	90,774	0,581
513	transistor	0,00613	0,061	90,835	1,163
514	troyano	0,00458	0,030	90,865	0,581
515	Tuenti	0,13948	1,062	91,927	20,349
516	Twitter	0,06047	0,455	92,382	8,721
517	Uma	0,00480	0,030	92,413	0,581
518	unidad central	0,00796	0,061	92,473	1,163
519	unidad de CD	0,00245	0,030	92,504	0,581
520	unidad de control	0,00416	0,030	92,534	0,581
521	unidad de disco	0,00270	0,030	92,564	0,581
522	UNIX	0,00480	0,030	92,595	0,581
523	URL	0,01529	0,121	92,716	2,326
524	USB	0,11299	0,910	93,627	17,442
525	usuario	0,01450	0,152	93,778	2,907
526	velocidad	0,01421	0,121	93,900	2,326
527	ventana	0,03775	0,303	94,203	5,814
528	ventilador	0,01313	0,152	94,355	2,907
529	vídeo	0,02779	0,243	94,598	4,651
530	vídeo musical	0,00283	0,030	94,628	0,581
531	videocámara	0,02231	0,182	94,810	3,488
532	videoconferencia	0,00687	0,061	94,871	1,163
533	videojuego	0,03013	0,212	95,083	4,070
534	videojuego online	0,00297	0,030	95,114	0,581
535	videollamada	0,00528	0,030	95,144	0,581
536	vínculo	0,00396	0,030	95,175	0,581
537	virtual	0,00213	0,030	95,205	0,581
538	virus	0,09896	0,880	96,085	16,860
539	visualizar	0,00327	0,030	96,115	0,581
540	Vodafone	0,00223	0,030	96,146	0,581
541	web	0,06627	0,486	96,631	9,302
542	web 2.0	0,00312	0,030	96,662	0,581
543	webcam	0,01752	0,182	96,844	3,488
544	WebQuest	0,00807	0,061	96,904	1,163
545	wifi	0,09710	0,850	97,754	16,279
546	wiki	0,03627	0,303	98,058	5,814

547	Wikipedia	0,04926	0,425	98,483	8,140
548	Windows	0,08399	0,637	99,120	12,209
549	Windows Movie Maker	0,00327	0,030	99,150	0,581
550	Windows Vista	0,00114	0,030	99,181	0,581
551	WordReference	0,00767	0,091	99,272	1,744
552	World Wide Web	0,00416	0,030	99,302	0,581
553	www	0,00628	0,061	99,363	1,163
554	Yahoo	0,02428	0,243	99,605	4,651
555	Youtube	0,04771	0,395	100,000	7,558

1.2. Listado de palabras en español ordenado según el índice de disponibilidad

1.2.1. Partes del cuerpo

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	ojo	0,67264	3,340	3,340	87,719
2	cabeza	0,65670	2,895	6,235	76,023
3	mano	0,60631	3,073	9,308	80,702
4	dedo	0,56195	3,162	12,469	83,041
5	pierna	0,55972	3,340	15,809	87,719
6	nariz	0,52369	3,051	18,860	80,117
7	pie	0,51059	3,140	22,000	82,456
8	boca	0,50902	2,917	24,916	76,608
9	oreja	0,50052	2,961	27,878	77,778
10	brazo	0,49081	2,717	30,595	71,345
11	pelo	0,39946	2,316	32,910	60,819
12	uña	0,39057	2,627	35,538	69,006
13	cuello	0,33162	2,160	37,698	56,725
14	hombro	0,30344	2,115	39,813	55,556
15	rodilla	0,30191	2,338	42,151	61,404
16	pecho	0,28416	2,093	44,244	54,971
17	codo	0,28046	1,937	46,181	50,877
18	espalda	0,27851	2,138	48,319	56,140
19	diente	0,27781	2,004	50,323	52,632
20	ceja	0,26571	1,781	52,104	46,784
21	pestaña	0,23095	1,648	53,752	43,275
22	tobillo	0,20883	1,759	55,511	46,199
23	cara	0,20682	1,136	56,647	29,825
24	corazón	0,20047	1,737	58,383	45,614
25	lengua	0,18389	1,291	59,675	33,918
26	muñeca	0,15826	1,202	60,877	31,579
27	labio	0,15121	1,002	61,879	26,316
28	estómago	0,15087	1,291	63,171	33,918
29	pulmón	0,14613	1,291	64,462	33,918
30	cadera	0,14499	1,202	65,665	31,579
31	barriga	0,13959	1,113	66,778	29,240
32	mejilla	0,12948	0,980	67,758	25,731
33	culo	0,12697	1,113	68,871	29,240
34	frente	0,11345	0,802	69,673	21,053
35	antebrazo	0,11029	0,802	70,474	21,053
36	riñón	0,10946	1,047	71,521	27,485
37	muslo	0,10519	0,824	72,345	21,637
38	cintura	0,10338	0,779	73,124	20,468
39	hígado	0,09071	0,824	73,948	21,637
40	tronco	0,09034	0,445	74,393	11,696

41	hueso	0,09022	0,735	75,128	19,298
42	barbilla	0,08991	0,757	75,885	19,883
43	cerebro	0,08533	0,646	76,531	16,959
44	ombbligo	0,08082	0,646	77,177	16,959
45	músculo	0,07375	0,579	77,756	15,205
46	pantorrilla	0,06157	0,557	78,312	14,620
47	intestino	0,05382	0,534	78,847	14,035
48	cráneo	0,05157	0,289	79,136	7,602
49	oído	0,05106	0,312	79,448	8,187
50	torso	0,04984	0,334	79,782	8,772
51	gemelo	0,04983	0,445	80,227	11,696
52	vena	0,04956	0,512	80,739	13,450
53	vientre	0,04519	0,334	81,073	8,772
54	tórax	0,04501	0,312	81,385	8,187
55	extremidad	0,04358	0,223	81,608	5,848
56	arteria	0,04306	0,445	82,053	11,696
57	columna (vertebral)	0,03959	0,379	82,432	9,942
58	dedo pulgar	0,03862	0,289	82,721	7,602
59	pelvis	0,03814	0,356	83,077	9,357
60	abdomen	0,03769	0,312	83,389	8,187
61	garganta	0,03753	0,379	83,768	9,942
62	clavícula	0,03685	0,289	84,057	7,602
63	pene	0,03668	0,289	84,346	7,602
64	páncreas	0,03662	0,334	84,680	8,772
65	mentón	0,03625	0,289	84,970	7,602
66	tímpano	0,03563	0,245	85,215	6,433
67	pómulo	0,03549	0,289	85,504	7,602
68	nudillo	0,03464	0,312	85,816	8,187
69	laringe	0,03418	0,334	86,150	8,772
70	bíceps	0,03371	0,312	86,462	8,187
71	piel	0,03288	0,334	86,796	8,772
72	nuca	0,03239	0,267	87,063	7,018
73	muela	0,03239	0,223	87,286	5,848
74	lóbulo	0,03167	0,245	87,531	6,433
75	fémur	0,03000	0,267	87,798	7,018
76	talón	0,02936	0,334	88,132	8,772
77	faringe	0,02840	0,312	88,444	8,187
78	cabello	0,02818	0,178	88,622	4,678
79	párpado	0,02800	0,267	88,889	7,018
80	vagina	0,02792	0,223	89,112	5,848
81	esófago	0,02708	0,289	89,401	7,602
82	sangre	0,02667	0,289	89,690	7,602
83	esternón	0,02570	0,267	89,958	7,018
84	vello	0,02487	0,223	90,180	5,848
85	teta	0,02282	0,200	90,381	5,263
86	cuádriceps	0,02172	0,223	90,603	5,848
87	palma de la mano	0,02123	0,134	90,737	3,509
88	nalga	0,02097	0,178	90,915	4,678
89	trasero	0,01996	0,200	91,116	5,263

90	tríceps	0,01994	0,178	91,294	4,678
91	bigote	0,01907	0,111	91,405	2,924
92	axila	0,01867	0,134	91,539	3,509
93	pupila	0,01837	0,178	91,717	4,678
94	entrecejo	0,01818	0,111	91,828	2,924
95	costilla	0,01774	0,200	92,029	5,263
96	bazo	0,01716	0,178	92,207	4,678
97	cuero cabelludo	0,01689	0,089	92,296	2,339
98	iris	0,01628	0,156	92,452	4,094
99	pezón	0,01608	0,134	92,585	3,509
100	tráquea	0,01545	0,223	92,808	5,848
101	falange	0,01516	0,134	92,941	3,509
102	barba	0,01504	0,111	93,053	2,924
103	meñique	0,01501	0,111	93,164	2,924
104	vejiga	0,01433	0,178	93,342	4,678
105	esternocleidomastoideo	0,01420	0,156	93,498	4,094
106	colmillo	0,01415	0,111	93,609	2,924
107	orificio nasal	0,01360	0,067	93,676	1,754
108	dedo índice	0,01248	0,111	93,788	2,924
109	radio	0,01228	0,134	93,921	3,509
110	nuez	0,01208	0,111	94,033	2,924
111	aparato reproductor	0,01204	0,111	94,144	2,924
112	esqueleto	0,01195	0,111	94,255	2,924
113	sien	0,01193	0,089	94,344	2,339
114	tibia	0,01163	0,111	94,456	2,924
115	ano	0,01151	0,111	94,567	2,924
116	aorta	0,01121	0,089	94,656	2,339
117	órgano	0,01102	0,111	94,767	2,924
118	michelín	0,01098	0,111	94,879	2,924
119	testículo	0,01071	0,089	94,968	2,339
120	cúbito	0,01071	0,111	95,079	2,924
121	peroné	0,01046	0,089	95,168	2,339
122	útero	0,01041	0,134	95,302	3,509
123	tendón	0,01033	0,067	95,369	1,754
124	cerebelo	0,01005	0,067	95,435	1,754
125	abdominal	0,00976	0,111	95,547	2,924
126	mandíbula	0,00964	0,089	95,636	2,339
127	glúteo	0,00960	0,089	95,725	2,339
128	ovario	0,00951	0,111	95,836	2,924
129	ingle	0,00936	0,067	95,903	1,754
130	planta del pie	0,00877	0,089	95,992	2,339
131	genital	0,00858	0,089	96,081	2,339
132	moflete	0,00848	0,089	96,170	2,339
133	escroto	0,00843	0,045	96,215	1,170
134	pubis	0,00831	0,089	96,304	2,339
135	aparato digestivo	0,00813	0,067	96,371	1,754
136	vulva	0,00789	0,045	96,415	1,170
137	espinas dorsal	0,00771	0,067	96,482	1,754
138	arruga	0,00749	0,067	96,549	1,754

139	glóbulo	0,00722	0,045	96,593	1,170
140	fosa nasal	0,00676	0,067	96,660	1,754
141	lunar	0,00658	0,089	96,749	2,339
142	intestino grueso	0,00629	0,089	96,838	2,339
143	capilar	0,00619	0,045	96,883	1,170
144	intestino delgado	0,00613	0,089	96,972	2,339
145	paleta	0,00602	0,045	97,016	1,170
146	jugo gástrico	0,00590	0,067	97,083	1,754
147	isquiotibial	0,00588	0,067	97,150	1,754
148	aparato excretor	0,00572	0,045	97,194	1,170
149	encia	0,00544	0,045	97,239	1,170
150	entropierna	0,00542	0,045	97,283	1,170
151	cuerpo	0,00528	0,022	97,306	0,585
152	cuerda vocal	0,00523	0,067	97,373	1,754
153	pectoral	0,00519	0,045	97,417	1,170
154	espinilla	0,00508	0,067	97,484	1,754
155	trompa de Falopio	0,00507	0,045	97,528	1,170
156	córnea	0,00505	0,045	97,573	1,170
157	apófisis	0,00502	0,022	97,595	0,585
158	glóbulo rojo	0,00491	0,045	97,640	1,170
159	apéndice	0,00491	0,045	97,684	1,170
160	uretra	0,00477	0,045	97,729	1,170
161	globo ocular	0,00477	0,022	97,751	0,585
162	omópato	0,00454	0,045	97,796	1,170
163	cavidad bucal	0,00453	0,022	97,818	0,585
164	seno	0,00453	0,022	97,840	0,585
165	dedo del pie	0,00445	0,045	97,885	1,170
166	tiroides	0,00445	0,045	97,929	1,170
167	saliva	0,00443	0,045	97,974	1,170
168	cabellera	0,00430	0,022	97,996	0,585
169	cogote	0,00430	0,022	98,018	0,585
170	articulación	0,00426	0,067	98,085	1,754
171	caries	0,00419	0,045	98,130	1,170
172	rostro	0,00415	0,045	98,174	1,170
173	ligamento	0,00414	0,045	98,219	1,170
174	médula espinal	0,00408	0,045	98,263	1,170
175	tripa	0,00405	0,045	98,308	1,170
176	tubo digestivo	0,00400	0,045	98,352	1,170
177	incisivo	0,00389	0,022	98,375	0,585
178	casco	0,00369	0,022	98,397	0,585
179	cerumen	0,00369	0,022	98,419	0,585
180	flequillo	0,00369	0,022	98,441	0,585
181	panza	0,00369	0,022	98,464	0,585
182	diente canino	0,00351	0,022	98,486	0,585
183	lumbar	0,00351	0,022	98,508	0,585
184	médula ósea	0,00351	0,022	98,530	0,585
185	nervio	0,00351	0,022	98,553	0,585
186	órgano sexual	0,00351	0,022	98,575	0,585
187	paladar	0,00351	0,022	98,597	0,585

188	cervical	0,00347	0,045	98,642	1,170
189	plexo	0,00345	0,045	98,686	1,170
190	empeine	0,00317	0,022	98,709	0,585
191	isquion	0,00317	0,022	98,731	0,585
192	úvula	0,00317	0,022	98,753	0,585
193	alveolo	0,00310	0,045	98,798	1,170
194	punto del pie	0,00301	0,022	98,820	0,585
195	sexo	0,00301	0,022	98,842	0,585
196	trompa de Eustaquio	0,00301	0,022	98,864	0,585
197	menisco	0,00293	0,045	98,909	1,170
198	vesícula	0,00280	0,045	98,953	1,170
199	alma	0,00272	0,022	98,976	0,585
200	lesión	0,00272	0,022	98,998	0,585
201	bronquio	0,00260	0,045	99,043	1,170
202	cardenal	0,00258	0,022	99,065	0,585
203	clitoris	0,00258	0,022	99,087	0,585
204	cutícula	0,00258	0,022	99,109	0,585
205	ganglio	0,00258	0,022	99,132	0,585
206	perilla	0,00258	0,022	99,154	0,585
207	puño	0,00258	0,022	99,176	0,585
208	zona lumbar	0,00258	0,022	99,198	0,585
209	huevo	0,00245	0,022	99,221	0,585
210	celulitis	0,00233	0,022	99,243	0,585
211	yugular	0,00233	0,022	99,265	0,585
212	colon	0,00221	0,022	99,287	0,585
213	dedo corazón	0,00221	0,022	99,310	0,585
214	bíceps femoral	0,00210	0,022	99,332	0,585
215	cachete	0,00210	0,022	99,354	0,585
216	deporte	0,00210	0,022	99,377	0,585
217	húmero	0,00210	0,022	99,399	0,585
218	recto	0,00210	0,022	99,421	0,585
219	condición física	0,00200	0,022	99,443	0,585
220	linfa	0,00200	0,022	99,466	0,585
221	viscera	0,00200	0,022	99,488	0,585
222	cava	0,00190	0,022	99,510	0,585
223	aparato fonador	0,00180	0,022	99,532	0,585
224	glóbulo blanco	0,00180	0,022	99,555	0,585
225	hipotálamo	0,00180	0,022	99,577	0,585
226	glándula	0,00172	0,022	99,599	0,585
227	hoyuelo	0,00172	0,022	99,621	0,585
228	masa muscular	0,00163	0,022	99,644	0,585
229	anular	0,00155	0,022	99,666	0,585
230	grano	0,00155	0,022	99,688	0,585
231	mucosa	0,00155	0,022	99,711	0,585
232	papada	0,00155	0,022	99,733	0,585
233	diente de leche	0,00147	0,022	99,755	0,585
234	célula	0,00140	0,022	99,777	0,585
235	coxis	0,00140	0,022	99,800	0,585
236	cristalino	0,00140	0,022	99,822	0,585

237	muela del juicio	0,00140	0,022	99,844	0,585
238	costado	0,00133	0,022	99,866	0,585
239	deltoides	0,00126	0,022	99,889	0,585
240	rótula	0,00108	0,022	99,911	0,585
241	óvulo	0,00093	0,022	99,933	0,585
242	bilis	0,00084	0,022	99,955	0,585
243	aurícula	0,00076	0,022	99,978	0,585
244	ventrículo	0,00072	0,022	100,000	0,585

1.2.2. La ropa

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	pantalón	0,80262	4,337	4,337	95,322
2	camiseta	0,61688	3,619	7,956	79,532
3	camisa	0,61591	3,619	11,575	79,532
4	falda	0,58022	3,539	15,114	77,778
5	jersey	0,56256	3,326	18,441	73,099
6	zapato	0,51552	3,619	22,060	79,532
7	chaqueta	0,49474	3,246	25,306	71,345
8	calcetín	0,47308	3,566	28,872	78,363
9	vestido	0,39388	2,555	31,426	56,140
10	bufanda	0,38735	2,821	34,247	61,988
11	abrigo	0,36013	2,608	36,855	57,310
12	media	0,29144	2,342	39,196	51,462
13	brega	0,27961	2,288	41,485	50,292
14	calzoncillo	0,26147	2,182	43,667	47,953
15	guante	0,25640	2,076	45,742	45,614
16	sujetador	0,23713	1,889	47,632	41,520
17	gorro	0,23632	1,836	49,468	40,351
18	(pantalón) vaquero	0,23517	1,650	51,118	36,257
19	bota	0,22251	1,969	53,087	43,275
20	sombrero	0,21252	1,543	54,630	33,918
21	rebeca	0,21181	1,517	56,147	33,333
22	corbata	0,19370	1,570	57,717	34,503
23	cinturón	0,18405	1,543	59,260	33,918
24	gorra	0,18242	1,384	60,644	30,409
25	blusa	0,16950	1,118	61,762	24,561
26	pañuelo	0,16440	1,384	63,145	30,409
27	chaleco	0,16244	1,330	64,476	29,240
28	chaquetón	0,15527	1,171	65,647	25,731
29	pijama	0,14543	1,251	66,897	27,485
30	tacón	0,13973	1,304	68,201	28,655
31	chándal	0,13884	1,224	69,425	26,901
32	zapatilla	0,13020	1,064	70,490	23,392
33	traje	0,12210	0,958	71,448	21,053
34	tanga	0,11956	1,064	72,512	23,392
35	sudadera	0,11847	0,958	73,470	21,053
36	bañador	0,11615	1,064	74,534	23,392
37	tenis	0,11214	0,958	75,492	21,053
38	sandalia	0,07576	0,665	76,158	14,620
39	camisón	0,07321	0,559	76,716	12,281
40	chancla	0,07167	0,665	77,382	14,620
41	ropa interior	0,07016	0,585	77,967	12,865
42	chubasquero	0,06810	0,585	78,552	12,865
43	minifalda	0,06550	0,612	79,164	13,450
44	polo	0,06084	0,426	79,590	9,357
45	zapatilla de deporte	0,06039	0,506	80,096	11,111

46	short	0,06001	0,532	80,628	11,696
47	anorak	0,05993	0,506	81,134	11,111
48	gabardina	0,05913	0,506	81,639	11,111
49	bolso	0,05500	0,532	82,171	11,696
50	bata	0,05385	0,506	82,677	11,111
51	cazadora	0,05236	0,399	83,076	8,772
52	biquini	0,05188	0,532	83,608	11,696
53	top	0,04848	0,426	84,034	9,357
54	pajarita	0,04704	0,452	84,486	9,942
55	forro polar	0,04613	0,399	84,886	8,772
56	leotardo	0,04435	0,399	85,285	8,772
57	fular	0,04350	0,399	85,684	8,772
58	pamela	0,03915	0,293	85,977	6,433
59	chal	0,03768	0,373	86,349	8,187
60	botín	0,03635	0,319	86,668	7,018
61	liguero	0,03482	0,293	86,961	6,433
62	traje de chaqueta	0,03367	0,293	87,254	6,433
63	collar	0,03302	0,293	87,547	6,433
64	levita	0,03178	0,239	87,786	5,263
65	pantalón corto	0,03023	0,239	88,026	5,263
66	bermudas	0,03010	0,213	88,238	4,678
67	peto	0,02958	0,293	88,531	6,433
68	boina	0,02955	0,239	88,771	5,263
69	plumón	0,02791	0,213	88,984	4,678
70	gafa	0,02669	0,293	89,276	6,433
71	tres cuartos	0,02650	0,186	89,462	4,094
72	camiseta interior	0,02625	0,266	89,729	5,848
73	pendiente	0,02448	0,239	89,968	5,263
74	manoletina	0,02422	0,239	90,208	5,263
75	esmoquin	0,02403	0,213	90,420	4,678
76	pulsera	0,02363	0,239	90,660	5,263
77	impermeable	0,02274	0,213	90,873	4,678
78	albornoz	0,02265	0,239	91,112	5,263
79	leggings	0,02245	0,239	91,352	5,263
80	suéter	0,02187	0,160	91,511	3,509
81	anillo	0,02183	0,213	91,724	4,678
82	panty	0,02040	0,160	91,884	3,509
83	correa	0,01935	0,186	92,070	4,094
84	mono	0,01900	0,160	92,230	3,509
85	bóxer	0,01849	0,186	92,416	4,094
86	gemelo	0,01818	0,160	92,576	3,509
87	manopla	0,01753	0,133	92,709	2,924
88	bailarina	0,01555	0,160	92,869	3,509
89	orejera	0,01541	0,133	93,002	2,924
90	mocasín	0,01527	0,133	93,135	2,924
91	cárdigan	0,01454	0,160	93,294	3,509
92	corpiño	0,01421	0,133	93,427	2,924
93	estola	0,01412	0,080	93,507	1,754
94	mallá	0,01372	0,133	93,640	2,924

95	blusón	0,01241	0,106	93,747	2,339
96	babucha	0,01161	0,080	93,827	1,754
97	slip	0,01156	0,133	93,960	2,924
98	botón	0,01136	0,106	94,066	2,339
99	tocado	0,01135	0,080	94,146	1,754
100	gafa de sol	0,01110	0,080	94,226	1,754
101	uniforme	0,01110	0,133	94,359	2,924
102	reloj	0,01103	0,106	94,465	2,339
103	tejano	0,01061	0,053	94,518	1,170
104	chamarreta	0,01035	0,080	94,598	1,754
105	pantalón de pinzas	0,01019	0,080	94,678	1,754
106	culote	0,00981	0,080	94,758	1,754
107	calzón	0,00978	0,053	94,811	1,170
108	tirante	0,00933	0,133	94,944	2,924
109	felpa	0,00922	0,080	95,024	1,754
110	alianza	0,00907	0,053	95,077	1,170
111	camisa de dormir	0,00870	0,080	95,157	1,754
112	bombín	0,00839	0,053	95,210	1,170
113	cremallera	0,00820	0,080	95,290	1,754
114	ropa	0,00818	0,053	95,343	1,170
115	vestido de novia	0,00814	0,053	95,396	1,170
116	americana	0,00808	0,080	95,476	1,754
117	pashmina	0,00798	0,106	95,583	2,339
118	(pantalón) pirata	0,00797	0,080	95,663	1,754
119	bota de agua	0,00792	0,080	95,742	1,754
120	pantalón de pana	0,00773	0,053	95,796	1,170
121	delantal	0,00768	0,080	95,875	1,754
122	capa	0,00756	0,080	95,955	1,754
123	túnica	0,00722	0,080	96,035	1,754
124	poncho	0,00670	0,080	96,115	1,754
125	falda pantalón	0,00653	0,053	96,168	1,170
126	bota alta	0,00652	0,053	96,221	1,170
127	seda	0,00652	0,053	96,275	1,170
128	algodón	0,00619	0,053	96,328	1,170
129	pantalón chino	0,00602	0,053	96,381	1,170
130	burka	0,00599	0,053	96,434	1,170
131	talla grande	0,00585	0,027	96,461	0,585
132	frac	0,00582	0,080	96,541	1,754
133	traje de baño	0,00575	0,053	96,594	1,170
134	body	0,00556	0,027	96,621	0,585
135	talla pequeña	0,00556	0,027	96,647	0,585
136	triquini	0,00544	0,053	96,700	1,170
137	velo	0,00539	0,053	96,754	1,170
138	colorido	0,00528	0,027	96,780	0,585
139	colgante	0,00520	0,080	96,860	1,754
140	cordón	0,00514	0,053	96,913	1,170
141	corsé	0,00508	0,053	96,966	1,170
142	pareo	0,00507	0,053	97,020	1,170
143	alpargata	0,00505	0,053	97,073	1,170

144	ropa deportiva	0,00503	0,053	97,126	1,170
145	diminuto	0,00502	0,027	97,153	0,585
146	manga a la sisa	0,00502	0,027	97,179	0,585
147	chilaba	0,00496	0,053	97,233	1,170
148	anteojo	0,00477	0,027	97,259	0,585
149	calzado	0,00477	0,027	97,286	0,585
150	niqui	0,00477	0,027	97,312	0,585
151	brazalete	0,00466	0,053	97,366	1,170
152	bolsillo	0,00458	0,053	97,419	1,170
153	gargantilla	0,00456	0,053	97,472	1,170
154	jeans	0,00453	0,027	97,499	0,585
155	torera	0,00453	0,027	97,525	0,585
156	turbante	0,00453	0,027	97,552	0,585
157	chaqué	0,00450	0,053	97,605	1,170
158	enagua	0,00432	0,053	97,658	1,170
159	camisola	0,00430	0,027	97,685	0,585
160	lencería	0,00409	0,027	97,712	0,585
161	palabra de honor	0,00409	0,027	97,738	0,585
162	toalla	0,00404	0,053	97,791	1,170
163	pasamontañas	0,00391	0,053	97,845	1,170
164	burkini	0,00389	0,027	97,871	0,585
165	falda de tubo	0,00389	0,027	97,898	0,585
166	capucha	0,00384	0,053	97,951	1,170
167	blazer	0,00369	0,027	97,978	0,585
168	falda de vuelo	0,00369	0,027	98,004	0,585
169	pantufla	0,00351	0,027	98,031	0,585
170	percha	0,00351	0,027	98,057	0,585
171	Borsalino	0,00333	0,027	98,084	0,585
172	bota de media caña	0,00333	0,027	98,111	0,585
173	bota de montaña	0,00333	0,027	98,137	0,585
174	zapatilla de casa	0,00333	0,027	98,164	0,585
175	mantón de Manila	0,00323	0,053	98,217	1,170
176	picardía	0,00320	0,053	98,270	1,170
177	conjunto	0,00317	0,027	98,297	0,585
178	joya	0,00317	0,027	98,324	0,585
179	pichi	0,00317	0,027	98,350	0,585
180	sombrero				
	chambergó	0,00317	0,027	98,377	0,585
181	faja	0,00310	0,053	98,430	1,170
182	bastón	0,00301	0,027	98,457	0,585
183	complemento	0,00301	0,027	98,483	0,585
184	manga	0,00301	0,027	98,510	0,585
185	vestido de noche	0,00301	0,027	98,536	0,585
186	Wonderbra	0,00301	0,027	98,563	0,585
187	disfraz	0,00289	0,053	98,616	1,170
188	chamarra	0,00286	0,027	98,643	0,585
189	hebilla	0,00286	0,027	98,670	0,585
190	vestido camisero	0,00286	0,027	98,696	0,585
191	lana	0,00272	0,027	98,723	0,585

192	reloj de pulsera	0,00272	0,027	98,749	0,585
193	calza	0,00258	0,027	98,776	0,585
194	pana	0,00258	0,027	98,803	0,585
195	mochila	0,00253	0,053	98,856	1,170
196	jersey de cuello alto	0,00245	0,027	98,882	0,585
197	mantón	0,00245	0,027	98,909	0,585
198	peluca	0,00245	0,027	98,936	0,585
199	equipación	0,00233	0,027	98,962	0,585
200	ropa blanca	0,00233	0,027	98,989	0,585
201	tienda	0,00233	0,027	99,015	0,585
202	bata de cola	0,00221	0,027	99,042	0,585
203	camiseta de tirantes	0,00221	0,027	99,069	0,585
204	probador	0,00221	0,027	99,095	0,585
205	ropa de color	0,00221	0,027	99,122	0,585
206	sostén	0,00221	0,027	99,148	0,585
207	forro	0,00210	0,027	99,175	0,585
208	vestido de flamenca	0,00210	0,027	99,202	0,585
209	accesorio	0,00200	0,027	99,228	0,585
210	formal	0,00200	0,027	99,255	0,585
211	media tupida	0,00200	0,027	99,282	0,585
212	vestido de tirantes	0,00200	0,027	99,308	0,585
213	deportivo	0,00190	0,027	99,335	0,585
214	informal	0,00190	0,027	99,361	0,585
215	traje de vestir	0,00190	0,027	99,388	0,585
216	moda	0,00180	0,027	99,415	0,585
217	traje de noche	0,00180	0,027	99,441	0,585
218	bisutería	0,00172	0,027	99,468	0,585
219	broche	0,00172	0,027	99,494	0,585
220	combinación	0,00172	0,027	99,521	0,585
221	estilo casual	0,00172	0,027	99,548	0,585
222	mantilla	0,00172	0,027	99,574	0,585
223	arreglarse	0,00163	0,027	99,601	0,585
224	sobrecamisa	0,00163	0,027	99,627	0,585
225	zanco	0,00163	0,027	99,654	0,585
226	antifaz	0,00155	0,027	99,681	0,585
227	braga (del cuello)	0,00155	0,027	99,707	0,585
228	zueco	0,00155	0,027	99,734	0,585
229	casaca	0,00147	0,027	99,761	0,585
230	muñequera	0,00147	0,027	99,787	0,585
231	sombrero de copa	0,00140	0,027	99,814	0,585
232	abalorio	0,00133	0,027	99,840	0,585
233	ejecutivo	0,00133	0,027	99,867	0,585
234	gomilla	0,00133	0,027	99,894	0,585
235	mitón	0,00133	0,027	99,920	0,585
236	pañuelo palestino	0,00133	0,027	99,947	0,585
237	fajín	0,00120	0,027	99,973	0,585
238	bolsa	0,00093	0,027	100,000	0,585

1.2.3. Comidas y bebidas

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	agua	0,44275	2,256	2,256	60,234
2	tomate	0,34845	2,278	4,535	60,819
3	carne	0,32667	1,731	6,265	46,199
4	zummo	0,32295	1,774	8,039	47,368
5	cerveza	0,32009	1,731	9,770	46,199
6	pescado	0,31706	1,862	11,632	49,708
7	vino	0,31504	1,752	13,384	46,784
8	patata	0,30074	1,884	15,268	50,292
9	pan	0,29139	1,709	16,977	45,614
10	leche	0,28277	1,555	18,532	41,520
11	lechuga	0,27775	1,818	20,350	48,538
12	arroz	0,24437	1,533	21,884	40,936
13	Coca-Cola	0,21271	1,051	22,935	28,070
14	café	0,20338	1,139	24,074	30,409
15	fruta	0,20306	1,161	25,235	30,994
16	naranja	0,19793	1,336	26,572	35,673
17	verdura	0,19704	1,227	27,798	32,749
18	ensalada	0,19421	1,183	28,981	31,579
19	manzana	0,18350	1,227	30,208	32,749
20	pollo	0,17838	1,139	31,347	30,409
21	huevo	0,17571	1,073	32,421	28,655
22	refresco	0,17408	0,942	33,363	25,146
23	pasta	0,17210	1,073	34,436	28,655
24	lenteja	0,15435	0,942	35,378	25,146
25	jamón	0,14388	1,117	36,495	29,825
26	chocolate	0,14068	0,854	37,349	22,807
27	zanahoria	0,13648	0,986	38,335	26,316
28	té	0,13456	0,789	39,124	21,053
29	pizza	0,13286	0,854	39,978	22,807
30	limón	0,13157	0,942	40,920	25,146
31	plátano	0,12699	0,920	41,840	24,561
32	cebolla	0,12478	0,964	42,804	25,731
33	tortilla	0,12264	0,767	43,571	20,468
34	sopa	0,11987	0,701	44,272	18,713
35	ternera	0,11876	0,811	45,082	21,637
36	paella	0,11741	0,657	45,739	17,544
37	fresa	0,11083	0,811	46,550	21,637
38	queso	0,11047	0,898	47,448	23,977
39	filete	0,11021	0,635	48,083	16,959
40	melón	0,10963	0,767	48,850	20,468
41	batido	0,10591	0,613	49,463	16,374
42	pimiento	0,10175	0,767	50,230	20,468
43	pera	0,10017	0,701	50,931	18,713
44	cerdo	0,09519	0,635	51,566	16,959
45	hamburguesa	0,09424	0,526	52,092	14,035

46	potaje	0,09398	0,548	52,640	14,620
47	macarrón	0,09237	0,570	53,209	15,205
48	espagueti	0,09201	0,504	53,713	13,450
49	cereal	0,09162	0,657	54,370	17,544
50	sal	0,09122	0,701	55,071	18,713
51	sandía	0,08688	0,657	55,728	17,544
52	legumbre	0,08494	0,526	56,254	14,035
53	atún	0,08261	0,635	56,889	16,959
54	mantequilla	0,08239	0,723	57,612	19,298
55	yogur	0,07681	0,657	58,269	17,544
56	maíz	0,07650	0,526	58,795	14,035
57	aceite	0,07552	0,635	59,430	16,959
58	puchero	0,07433	0,416	59,847	11,111
59	güisqui	0,07139	0,416	60,263	11,111
60	azúcar	0,06724	0,460	60,723	12,281
61	Fanta	0,06550	0,329	61,051	8,772
62	piña	0,06522	0,526	61,577	14,035
63	melocotón	0,06460	0,504	62,081	13,450
64	chorizo	0,06287	0,438	62,519	11,696
65	bocadillo	0,06277	0,394	62,913	10,526
66	tarta	0,06275	0,438	63,352	11,696
67	ajo	0,06149	0,504	63,855	13,450
68	galleta	0,06106	0,416	64,272	11,111
69	gazpacho	0,05711	0,329	64,600	8,772
70	guisante	0,05544	0,416	65,016	11,111
71	pimienta	0,05350	0,416	65,433	11,111
72	coliflor	0,05276	0,394	65,827	10,526
73	helado	0,05161	0,394	66,221	10,526
74	tortilla de patatas	0,05125	0,263	66,484	7,018
75	mandarina	0,05121	0,394	66,878	10,526
76	pepino	0,05119	0,372	67,251	9,942
77	cordero	0,05073	0,307	67,558	8,187
78	ron	0,05032	0,285	67,842	7,602
79	calabacín	0,04988	0,394	68,237	10,526
80	judía	0,04899	0,350	68,587	9,357
81	salchicha	0,04889	0,350	68,938	9,357
82	garbanzo	0,04829	0,372	69,310	9,942
83	mermelada	0,04588	0,350	69,660	9,357
84	solomillo	0,04586	0,285	69,945	7,602
85	calabaza	0,04568	0,285	70,230	7,602
86	col	0,04469	0,307	70,537	8,187
87	berenjena	0,04418	0,263	70,800	7,018
88	cocido	0,04249	0,219	71,019	5,848
89	patata frita	0,04131	0,241	71,260	6,433
90	chuleta	0,04063	0,263	71,522	7,018
91	boquerón	0,04061	0,263	71,785	7,018
92	uva	0,04029	0,350	72,136	9,357
93	limonada	0,03992	0,263	72,399	7,018
94	lasaña	0,03895	0,241	72,640	6,433

95	haba	0,03826	0,263	72,903	7,018
96	pastel	0,03758	0,307	73,209	8,187
97	estofado	0,03752	0,219	73,428	5,848
98	pavo	0,03669	0,263	73,691	7,018
99	salsa	0,03567	0,263	73,954	7,018
100	fideo	0,03531	0,263	74,217	7,018
101	sándwich	0,03438	0,219	74,436	5,848
102	aceituna	0,03406	0,263	74,699	7,018
103	postre	0,03404	0,263	74,962	7,018
104	alcachofa	0,03324	0,219	75,181	5,848
105	calamar	0,03181	0,219	75,400	5,848
106	vinagre	0,03181	0,285	75,685	7,602
107	alubia	0,03143	0,197	75,882	5,263
108	gamba	0,03141	0,241	76,123	6,433
109	salmón	0,03050	0,197	76,320	5,263
110	morcilla	0,02962	0,175	76,495	4,678
111	infusión	0,02953	0,219	76,714	5,848
112	marisco	0,02858	0,219	76,933	5,848
113	cereza	0,02852	0,219	77,152	5,848
114	salchichón	0,02773	0,307	77,459	8,187
115	remolacha	0,02769	0,219	77,678	5,848
116	espinaca	0,02694	0,197	77,875	5,263
117	fabada	0,02672	0,175	78,050	4,678
118	mango	0,02580	0,197	78,248	5,263
119	harina	0,02550	0,197	78,445	5,263
120	dulce	0,02509	0,219	78,664	5,848
121	bizcocho	0,02502	0,197	78,861	5,263
122	nata	0,02476	0,197	79,058	5,263
123	albóndiga	0,02453	0,175	79,233	4,678
124	Cola Cao	0,02398	0,175	79,409	4,678
125	tinto de verano	0,02391	0,131	79,540	3,509
126	bacalao	0,02367	0,175	79,715	4,678
127	aguacate	0,02344	0,175	79,890	4,678
128	anchoa	0,02327	0,175	80,066	4,678
129	mejillón	0,02312	0,175	80,241	4,678
130	alcohol	0,02281	0,131	80,372	3,509
131	embutido	0,02191	0,175	80,548	4,678
132	champán	0,02181	0,131	80,679	3,509
133	jamón york	0,02178	0,131	80,811	3,509
134	apio	0,02171	0,175	80,986	4,678
135	granizada	0,02167	0,131	81,117	3,509
136	tónica	0,02122	0,153	81,271	4,094
137	vino tinto	0,02058	0,110	81,380	2,924
138	guiso	0,02047	0,088	81,468	2,339
139	perejil	0,02043	0,175	81,643	4,678
140	licor	0,02035	0,153	81,796	4,094
141	mayonesa	0,01999	0,197	81,993	5,263
142	croqueta	0,01982	0,131	82,125	3,509
143	chóped	0,01966	0,153	82,278	4,094

144	espárrago	0,01946	0,110	82,388	2,924
145	habichuela	0,01895	0,153	82,541	4,094
146	gaseosa	0,01894	0,131	82,673	3,509
147	chanquete	0,01887	0,131	82,804	3,509
148	almeja	0,01837	0,131	82,935	3,509
149	kiwi	0,01835	0,175	83,111	4,678
150	vodka	0,01831	0,131	83,242	3,509
151	tomate frito	0,01820	0,153	83,395	4,094
152	salmorejo	0,01810	0,131	83,527	3,509
153	lomo	0,01808	0,131	83,658	3,509
154	acelga	0,01795	0,131	83,790	3,509
155	chirimoya	0,01791	0,131	83,921	3,509
156	vegetal	0,01759	0,088	84,009	2,339
157	ketchup	0,01751	0,175	84,184	4,678
158	costilla	0,01745	0,110	84,294	2,924
159	almendra	0,01708	0,131	84,425	3,509
160	Bacardi	0,01674	0,066	84,491	1,754
161	frambuesa	0,01653	0,131	84,622	3,509
162	gin-tonic	0,01651	0,088	84,710	2,339
163	Seven Up	0,01635	0,088	84,797	2,339
164	lenguado	0,01619	0,088	84,885	2,339
165	miel	0,01608	0,110	84,995	2,924
166	Nestea	0,01605	0,088	85,082	2,339
167	mora	0,01597	0,110	85,192	2,924
168	margarina	0,01594	0,153	85,345	4,094
169	manzanilla	0,01564	0,088	85,433	2,339
170	hortaliza	0,01545	0,131	85,564	3,509
171	champiñón	0,01536	0,131	85,696	3,509
172	tila	0,01536	0,088	85,783	2,339
173	paté	0,01506	0,131	85,915	3,509
174	nuez	0,01490	0,175	86,090	4,678
175	sardina	0,01465	0,131	86,221	3,509
176	panini	0,01461	0,088	86,309	2,339
177	natilla	0,01449	0,131	86,440	3,509
178	cava	0,01443	0,088	86,528	2,339
179	pato	0,01414	0,088	86,616	2,339
180	entrecot	0,01398	0,110	86,725	2,924
181	magdalena	0,01397	0,131	86,857	3,509
182	soja	0,01378	0,110	86,966	2,924
183	flan	0,01373	0,110	87,076	2,924
184	puré	0,01359	0,153	87,229	4,094
185	albaricoque	0,01356	0,110	87,338	2,924
186	empanada	0,01352	0,088	87,426	2,339
187	menta poleo	0,01338	0,088	87,514	2,339
188	tortilla francesa	0,01330	0,066	87,579	1,754
189	jamón serrano	0,01324	0,088	87,667	2,339
190	cruasán	0,01306	0,088	87,755	2,339
191	arroz con leche	0,01301	0,110	87,864	2,924
192	orégano	0,01290	0,110	87,974	2,924

193	conejo	0,01280	0,088	88,061	2,339
194	ginebra	0,01267	0,088	88,149	2,339
195	canela	0,01242	0,110	88,258	2,924
196	seta	0,01190	0,088	88,346	2,339
197	fondue	0,01173	0,066	88,412	1,754
198	vino blanco	0,01156	0,066	88,478	1,754
199	Pepsi	0,01142	0,044	88,521	1,170
200	caldo	0,01131	0,088	88,609	2,339
201	ensaladilla rusa	0,01114	0,110	88,719	2,924
202	rape	0,01107	0,066	88,784	1,754
203	trucha	0,01106	0,066	88,850	1,754
204	pollo asado	0,01104	0,066	88,916	1,754
205	crepe	0,01103	0,066	88,981	1,754
206	beicon	0,01097	0,131	89,113	3,509
207	porra	0,01097	0,066	89,179	1,754
208	crema	0,01094	0,088	89,266	2,339
209	ajo blanco	0,01068	0,088	89,354	2,339
210	nectarina	0,01061	0,088	89,441	2,339
211	tostada	0,01052	0,088	89,529	2,339
212	JB	0,01038	0,044	89,573	1,170
213	puerro	0,01036	0,088	89,660	2,339
214	fruto seco	0,01002	0,066	89,726	1,754
215	foie gras	0,00996	0,088	89,814	2,339
216	castaña	0,00995	0,088	89,901	2,339
217	Aquarius	0,00992	0,044	89,945	1,170
218	aguardiente	0,00989	0,044	89,989	1,170
219	Cacique	0,00989	0,044	90,033	1,170
220	bebida isotónica	0,00970	0,044	90,077	1,170
221	pez espada	0,00968	0,088	90,164	2,339
222	sushi	0,00967	0,044	90,208	1,170
223	sangría	0,00966	0,066	90,274	1,754
224	especia	0,00962	0,088	90,361	2,339
225	calamarito	0,00957	0,066	90,427	1,754
226	gazpachuelo	0,00956	0,044	90,471	1,170
227	cóctel	0,00943	0,044	90,515	1,170
228	buey	0,00941	0,066	90,581	1,754
229	rosada	0,00927	0,066	90,646	1,754
230	nube	0,00920	0,044	90,690	1,170
231	níspero	0,00912	0,088	90,778	2,339
232	chicle	0,00899	0,044	90,821	1,170
233	trigo	0,00895	0,066	90,887	1,754
234	zum de naranja	0,00889	0,088	90,975	2,339
235	coñac	0,00886	0,044	91,019	1,170
236	tocino	0,00886	0,044	91,062	1,170
237	callos	0,00869	0,044	91,106	1,170
238	cebada	0,00863	0,066	91,172	1,754
239	ave	0,00857	0,044	91,216	1,170
240	limoncello	0,00857	0,044	91,260	1,170
241	flamenquín	0,00856	0,066	91,325	1,754

242	Red Bull	0,00852	0,044	91,369	1,170
243	repollo	0,00845	0,066	91,435	1,754
244	patata asada	0,00829	0,066	91,501	1,754
245	arroz a la cubana	0,00829	0,066	91,566	1,754
246	picatoste	0,00818	0,066	91,632	1,754
247	pepinillo	0,00817	0,088	91,720	2,339
248	desayuno	0,00816	0,044	91,763	1,170
249	perrito (caliente)	0,00816	0,044	91,807	1,170
250	migas	0,00815	0,088	91,895	2,339
251	mortadela	0,00803	0,088	91,982	2,339
252	porra antequerana	0,00795	0,066	92,048	1,754
253	pero	0,00790	0,044	92,092	1,170
254	caramelo	0,00784	0,066	92,158	1,754
255	banana	0,00778	0,044	92,202	1,170
256	brioche	0,00778	0,044	92,245	1,170
257	salsa rosa	0,00770	0,066	92,311	1,754
258	granada	0,00765	0,066	92,377	1,754
259	merluza	0,00762	0,066	92,442	1,754
260	manteca	0,00761	0,044	92,486	1,170
261	mortadela de aceitunas	0,00756	0,066	92,552	1,754
262	batata	0,00752	0,066	92,618	1,754
263	tofu	0,00748	0,066	92,683	1,754
264	avena	0,00718	0,044	92,727	1,170
265	salsa de yogur	0,00702	0,066	92,793	1,754
266	Sprite	0,00698	0,044	92,837	1,170
267	canelón	0,00686	0,066	92,903	1,754
268	leche condensada	0,00679	0,044	92,946	1,170
269	pitufo	0,00679	0,044	92,990	1,170
270	boquerón en vinagre	0,00676	0,044	93,034	1,170
271	vino rosado	0,00659	0,044	93,078	1,170
272	mero	0,00645	0,044	93,122	1,170
273	anís	0,00645	0,044	93,165	1,170
274	caballa	0,00628	0,044	93,209	1,170
275	brócoli	0,00627	0,044	93,253	1,170
276	azafrán	0,00627	0,044	93,297	1,170
277	gominola	0,00621	0,044	93,341	1,170
278	leche merengada	0,00613	0,044	93,384	1,170
279	chuchería	0,00603	0,044	93,428	1,170
280	Doritos	0,00585	0,022	93,450	0,585
281	judía verde	0,00584	0,044	93,494	1,170
282	napolitana	0,00576	0,044	93,538	1,170
283	bollo	0,00559	0,066	93,604	1,754
284	cazuela	0,00557	0,022	93,625	0,585
285	Dyc	0,00557	0,022	93,647	0,585

286	rábano	0,00557	0,044	93,691	1,170
287	bicarbonato	0,00556	0,044	93,735	1,170
288	ciruela	0,00556	0,088	93,823	2,339
289	zum de fruta	0,00553	0,044	93,866	1,170
290	sanjacobo	0,00547	0,066	93,932	1,754
291	lomo en manteca	0,00546	0,044	93,976	1,170
292	pionono	0,00534	0,044	94,020	1,170
293	pan de molde	0,00533	0,044	94,064	1,170
294	emblanco	0,00532	0,044	94,107	1,170
295	langostino	0,00532	0,044	94,151	1,170
296	pulpo	0,00532	0,044	94,195	1,170
297	avellana	0,00531	0,022	94,217	0,585
298	carne de cerdo	0,00531	0,022	94,239	0,585
299	sauvignon	0,00531	0,022	94,261	0,585
300	solomillo a la pimienta	0,00531	0,022	94,283	0,585
301	jamón ibérico	0,00531	0,044	94,326	1,170
302	tomillo	0,00531	0,044	94,370	1,170
303	cuscús	0,00513	0,044	94,414	1,170
304	pan integral	0,00513	0,044	94,458	1,170
305	Chardonnay	0,00506	0,022	94,480	0,585
306	entrante	0,00506	0,022	94,502	0,585
307	Fanta Naranja	0,00506	0,022	94,524	0,585
308	pinchito	0,00506	0,022	94,545	0,585
309	tarta de chocolate	0,00506	0,022	94,567	0,585
310	pisto	0,00489	0,044	94,611	1,170
311	hígado	0,00483	0,044	94,655	1,170
312	albariño	0,00483	0,022	94,677	0,585
313	brandy	0,00483	0,022	94,699	0,585
314	caldo gallego	0,00483	0,022	94,721	0,585
315	comida basura	0,00483	0,022	94,743	0,585
316	Fanta Limón	0,00483	0,022	94,765	0,585
317	fusilli	0,00483	0,022	94,786	0,585
318	pacharán	0,00483	0,022	94,808	0,585
319	palmera	0,00483	0,022	94,830	0,585
320	sake	0,00483	0,022	94,852	0,585
321	seso asado	0,00483	0,022	94,874	0,585
322	lima	0,00482	0,044	94,918	1,170
323	empanadilla	0,00482	0,044	94,962	1,170
324	cacao	0,00478	0,044	95,005	1,170
325	bebida alcohólica	0,00473	0,044	95,049	1,170
326	papaya	0,00473	0,044	95,093	1,170
327	pomelo	0,00468	0,044	95,137	1,170
328	puré de patata	0,00461	0,044	95,181	1,170
329	criadilla	0,00460	0,022	95,203	0,585
330	jibia	0,00460	0,022	95,225	0,585

331	kebab	0,00460	0,022	95,246	0,585
332	McDonald's	0,00460	0,022	95,268	0,585
333	mijo	0,00460	0,022	95,290	0,585
334	Ribera del Duero	0,00460	0,022	95,312	0,585
335	chistorra	0,00451	0,044	95,356	1,170
336	coco	0,00449	0,044	95,400	1,170
337	chupachús	0,00439	0,022	95,422	0,585
338	concha fina	0,00439	0,022	95,444	0,585
339	manita de cerdo	0,00439	0,022	95,465	0,585
340	rioja	0,00439	0,022	95,487	0,585
341	shawarma	0,00439	0,022	95,509	0,585
342	tintorera	0,00439	0,022	95,531	0,585
343	mojito	0,00418	0,022	95,553	0,585
344	pan cateto	0,00418	0,022	95,575	0,585
345	rebujito	0,00418	0,022	95,597	0,585
346	sofrito	0,00418	0,022	95,619	0,585
347	té verde	0,00418	0,022	95,641	0,585
348	tortellini	0,00418	0,022	95,663	0,585
349	hierbabuena	0,00407	0,044	95,706	1,170
350	roquefort	0,00404	0,044	95,750	1,170
351	brownie	0,00401	0,044	95,794	1,170
352	salami	0,00400	0,044	95,838	1,170
353	agua con gas	0,00399	0,022	95,860	0,585
354	baguette	0,00399	0,022	95,882	0,585
355	cardo	0,00399	0,022	95,904	0,585
356	chupito	0,00399	0,022	95,926	0,585
357	coquina	0,00399	0,022	95,947	0,585
358	cubalibre	0,00399	0,022	95,969	0,585
359	escarola	0,00399	0,022	95,991	0,585
360	pan de ajo	0,00399	0,022	96,013	0,585
361	potaje de acelgas	0,00399	0,022	96,035	0,585
362	sopa de picadillo	0,00399	0,022	96,057	0,585
363	membrillo	0,00390	0,044	96,101	1,170
364	panceta	0,00385	0,044	96,145	1,170
365	agua mineral	0,00380	0,022	96,166	0,585
366	brécol	0,00380	0,022	96,188	0,585
367	caldo Avecrem	0,00380	0,022	96,210	0,585
368	chipirón	0,00380	0,022	96,232	0,585
369	jabalí	0,00380	0,022	96,254	0,585
370	Sex on the Beach	0,00380	0,022	96,276	0,585
371	wrap	0,00380	0,022	96,298	0,585
372	queso azul	0,00377	0,044	96,342	1,170
373	fiambre	0,00374	0,044	96,386	1,170
374	golosina	0,00371	0,044	96,429	1,170
375	caqui	0,00364	0,044	96,473	1,170
376	camarón	0,00362	0,022	96,495	0,585

377	chacina	0,00362	0,022	96,517	0,585
378	fréjol	0,00362	0,022	96,539	0,585
379	horchata	0,00362	0,022	96,561	0,585
380	Martini	0,00362	0,022	96,583	0,585
381	pechuga	0,00362	0,022	96,605	0,585
382	ponche	0,00362	0,022	96,627	0,585
383	salsa boloñesa	0,00362	0,022	96,648	0,585
384	sidra	0,00362	0,022	96,670	0,585
385	valeriana	0,00362	0,022	96,692	0,585
386	mostaza	0,00346	0,044	96,736	1,170
387	azúcar moreno	0,00345	0,022	96,758	0,585
388	barra de pan	0,00345	0,022	96,780	0,585
389	carne a la plancha	0,00345	0,022	96,802	0,585
390	cigala	0,00345	0,022	96,824	0,585
391	espeto	0,00345	0,022	96,846	0,585
392	mitad	0,00345	0,022	96,867	0,585
393	pipa	0,00345	0,044	96,911	1,170
394	pimentón	0,00331	0,044	96,955	1,170
395	dorada	0,00329	0,044	96,999	1,170
396	berberecho	0,00329	0,022	97,021	0,585
397	huevo cocido	0,00329	0,022	97,043	0,585
398	mozzarella	0,00329	0,022	97,065	0,585
399	nugget	0,00329	0,022	97,087	0,585
400	pollo al curry	0,00329	0,022	97,108	0,585
401	revuelto de verdura	0,00329	0,022	97,130	0,585
402	sacarina	0,00329	0,022	97,152	0,585
403	vaca	0,00329	0,022	97,174	0,585
404	cebollita	0,00315	0,044	97,218	1,170
405	zum de piña	0,00314	0,044	97,262	1,170
406	bebida	0,00314	0,022	97,284	0,585
407	conserva vegetal	0,00314	0,022	97,306	0,585
408	Danone	0,00314	0,022	97,327	0,585
409	huevo de codorniz	0,00314	0,022	97,349	0,585
410	nabo	0,00314	0,022	97,371	0,585
411	potaje de lentejas	0,00314	0,022	97,393	0,585
412	cazuela de fideos	0,00299	0,022	97,415	0,585
413	conserva animal	0,00299	0,022	97,437	0,585
414	fideuá	0,00299	0,022	97,459	0,585
415	gamba rebozada	0,00299	0,022	97,481	0,585
416	mollete	0,00299	0,022	97,503	0,585
417	muslo	0,00299	0,022	97,525	0,585
418	palomita	0,00299	0,022	97,547	0,585
419	pimiento morrón	0,00299	0,022	97,568	0,585
420	soda	0,00299	0,022	97,590	0,585
421	aceituna rellena	0,00285	0,022	97,612	0,585
422	hojaldre	0,00285	0,022	97,634	0,585

423	jamón cocido	0,00285	0,022	97,656	0,585
424	naranjada	0,00285	0,022	97,678	0,585
425	sobrasada	0,00285	0,022	97,700	0,585
426	vainilla	0,00285	0,022	97,722	0,585
427	carpaccio	0,00272	0,022	97,744	0,585
428	caviar	0,00272	0,022	97,766	0,585
429	cena	0,00272	0,022	97,788	0,585
430	flan de huevo	0,00272	0,022	97,809	0,585
431	macedonia	0,00272	0,022	97,831	0,585
432	queso manchego	0,00272	0,022	97,853	0,585
433	almuerzo	0,00259	0,022	97,875	0,585
434	cocinar	0,00259	0,022	97,897	0,585
435	crema pastelera	0,00259	0,022	97,919	0,585
436	guarnición	0,00259	0,022	97,941	0,585
437	Puleva	0,00259	0,022	97,963	0,585
438	anguila	0,00247	0,022	97,985	0,585
439	boca de mar	0,00247	0,022	98,007	0,585
440	bombón	0,00247	0,022	98,028	0,585
441	calzone	0,00247	0,022	98,050	0,585
442	hambre	0,00247	0,022	98,072	0,585
443	riñón al jerez	0,00247	0,022	98,094	0,585
444	rollito de primavera	0,00247	0,022	98,116	0,585
445	absenta	0,00235	0,022	98,138	0,585
446	conejo a la cazadora	0,00235	0,022	98,160	0,585
447	fruto rojo	0,00235	0,022	98,182	0,585
448	gula	0,00235	0,022	98,204	0,585
449	maracuyá	0,00235	0,022	98,226	0,585
450	mate	0,00235	0,022	98,248	0,585
451	producto biológico	0,00235	0,022	98,269	0,585
452	ruca	0,00235	0,022	98,291	0,585
453	tarta de fruta	0,00235	0,022	98,313	0,585
454	ajeno	0,00224	0,022	98,335	0,585
455	burrito	0,00224	0,022	98,357	0,585
456	cangrejo	0,00224	0,022	98,379	0,585
457	gallina	0,00224	0,022	98,401	0,585
458	gallo	0,00224	0,022	98,423	0,585
459	patata a lo pobre	0,00224	0,022	98,445	0,585
460	tomate cherry	0,00224	0,022	98,467	0,585
461	alita de pollo	0,00214	0,022	98,489	0,585
462	arándano	0,00214	0,022	98,510	0,585
463	berenjena con miel	0,00214	0,022	98,532	0,585
464	carne mechada	0,00214	0,022	98,554	0,585
465	filete tártaro	0,00214	0,022	98,576	0,585
466	jengibre	0,00214	0,022	98,598	0,585
467	láudano	0,00214	0,022	98,620	0,585
468	lechal	0,00214	0,022	98,642	0,585

469	pasa	0,00214	0,022	98,664	0,585
470	revuelto	0,00214	0,022	98,686	0,585
471	anacardo	0,00204	0,022	98,708	0,585
472	choco	0,00204	0,022	98,729	0,585
473	churro	0,00204	0,022	98,751	0,585
474	gelatina	0,00204	0,022	98,773	0,585
475	revuelto de espárragos	0,00204	0,022	98,795	0,585
476	cacahuete	0,00194	0,022	98,817	0,585
477	chopo	0,00194	0,022	98,839	0,585
478	ingrediente	0,00194	0,022	98,861	0,585
479	La Casera	0,00194	0,022	98,883	0,585
480	ostra	0,00194	0,022	98,905	0,585
481	picadillo	0,00194	0,022	98,927	0,585
482	restaurante	0,00194	0,022	98,949	0,585
483	secreto ibérico	0,00194	0,022	98,970	0,585
484	zurrapa	0,00194	0,022	98,992	0,585
485	bistec	0,00185	0,022	99,014	0,585
486	mantecado	0,00185	0,022	99,036	0,585
487	muffin	0,00185	0,022	99,058	0,585
488	pipa de girasol	0,00185	0,022	99,080	0,585
489	piquito	0,00185	0,022	99,102	0,585
490	receta	0,00185	0,022	99,124	0,585
491	sirope	0,00185	0,022	99,146	0,585
492	tocino de cielo	0,00185	0,022	99,168	0,585
493	ciervo	0,00177	0,022	99,189	0,585
494	palmito	0,00177	0,022	99,211	0,585
495	pipa de calabaza	0,00177	0,022	99,233	0,585
496	rabo de toro	0,00177	0,022	99,255	0,585
497	rebanada	0,00177	0,022	99,277	0,585
498	sopa de marisco	0,00177	0,022	99,299	0,585
499	tapa	0,00177	0,022	99,321	0,585
500	zumو de melocotón	0,00177	0,022	99,343	0,585
501	alioli	0,00168	0,022	99,365	0,585
502	caña	0,00168	0,022	99,387	0,585
503	Kiko	0,00168	0,022	99,409	0,585
504	quesadilla	0,00168	0,022	99,430	0,585
505	queso de untar	0,00168	0,022	99,452	0,585
506	repostería	0,00168	0,022	99,474	0,585
507	gofre	0,00160	0,022	99,496	0,585
508	hielo	0,00160	0,022	99,518	0,585
509	melocotón en almíbar	0,00160	0,022	99,540	0,585
510	queso blanco	0,00160	0,022	99,562	0,585
511	rollo	0,00160	0,022	99,584	0,585
512	zumو de arándano	0,00160	0,022	99,606	0,585

513	ensalada de arroz	0,00153	0,022	99,628	0,585
514	papilla	0,00153	0,022	99,650	0,585
515	cuajada	0,00146	0,022	99,671	0,585
516	queso brie	0,00146	0,022	99,693	0,585
517	salmonete	0,00146	0,022	99,715	0,585
518	brote de soja	0,00132	0,022	99,737	0,585
519	cochinillo	0,00132	0,022	99,759	0,585
520	aperitivo	0,00126	0,022	99,781	0,585
521	condimento	0,00126	0,022	99,803	0,585
522	piñón	0,00126	0,022	99,825	0,585
523	salteado	0,00126	0,022	99,847	0,585
524	canapé	0,00120	0,022	99,869	0,585
525	ñoqui	0,00120	0,022	99,890	0,585
526	asado	0,00109	0,022	99,912	0,585
527	mantel	0,00071	0,022	99,934	0,585
528	cuchara	0,00068	0,022	99,956	0,585
529	cuchillo	0,00065	0,022	99,978	0,585
530	tenedor	0,00062	0,022	100,000	0,585

1.2.4. La escuela (muebles y materiales)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	pizarra	0,75097	4,061	4,061	88,304
2	silla	0,68505	3,792	7,854	82,456
3	mesa	0,66115	3,470	11,323	75,439
4	lápiz	0,62040	4,008	15,331	87,135
5	boli(grafo)	0,58391	3,846	19,177	83,626
6	libro	0,46042	3,416	22,593	74,269
7	tiza	0,45688	2,797	25,390	60,819
8	goma (de borrar)	0,41416	2,932	28,322	63,743
9	pupitre	0,41123	2,205	30,527	47,953
10	ordenador	0,35321	2,582	33,109	56,140
11	libreta	0,29204	2,152	35,261	46,784
12	regla	0,27437	2,071	37,332	45,029
13	estuche	0,27066	2,205	39,537	47,953
14	borrador	0,26026	1,694	41,232	36,842
15	rotulador	0,24592	1,910	43,141	41,520
16	cuaderno	0,23802	1,802	44,944	39,181
17	profesor	0,22485	1,399	46,342	30,409
18	mochila	0,22055	1,910	48,252	41,520
19	proyector	0,19520	1,399	49,650	30,409
20	folio	0,16718	1,452	51,103	31,579
21	alumno	0,16535	1,076	52,179	23,392
22	carpeta	0,16035	1,479	53,658	32,164
23	ventana	0,15467	1,183	54,841	25,731
24	sacapuntas	0,15239	1,130	55,971	24,561
25	puerta	0,15163	1,076	57,047	23,392
26	armario	0,14676	1,049	58,096	22,807
27	compás	0,14419	1,291	59,387	28,070
28	clase	0,11999	0,807	60,194	17,544
29	papel	0,11556	0,941	61,135	20,468
30	patio (de recreo)	0,11472	0,914	62,049	19,883
31	escuadra	0,10981	0,888	62,937	19,298
32	cartabón	0,10464	0,861	63,798	18,713
33	calculadora	0,08936	0,807	64,605	17,544
34	escritorio	0,08527	0,457	65,062	9,942
35	estantería	0,08346	0,592	65,654	12,865
36	biblioteca	0,07941	0,699	66,353	15,205
37	aula	0,07924	0,565	66,918	12,281
38	perchero	0,07641	0,619	67,536	13,450
39	papelera	0,07313	0,646	68,182	14,035
40	diccionario	0,06753	0,646	68,827	14,035
41	agenda (escolar)	0,06635	0,511	69,338	11,111
42	lápiz de color	0,06411	0,511	69,849	11,111
43	Tipp-Ex	0,06383	0,511	70,360	11,111
44	gimnasio	0,06358	0,511	70,871	11,111

45	recreo	0,06301	0,538	71,409	11,696
46	lápiz de cera	0,06255	0,538	71,947	11,696
47	tijera	0,06003	0,565	72,512	12,281
48	pegamento	0,05809	0,511	73,023	11,111
49	pantalla	0,05684	0,430	73,453	9,357
50	cañón	0,05671	0,430	73,884	9,357
51	pizarra digital	0,05500	0,323	74,207	7,018
52	portaminas	0,05487	0,430	74,637	9,357
53	cartulina	0,05400	0,565	75,202	12,281
54	mapa	0,05169	0,484	75,686	10,526
55	(ordenador) portátil	0,05058	0,350	76,036	7,602
56	pluma	0,05029	0,350	76,385	7,602
57	maleta	0,04922	0,377	76,762	8,187
58	tarima	0,04675	0,323	77,084	7,018
59	banco	0,04444	0,323	77,407	7,018
60	lámpara	0,04037	0,296	77,703	6,433
61	color	0,03888	0,377	78,080	8,187
62	grapadora	0,03779	0,350	78,429	7,602
63	examen	0,03711	0,350	78,779	7,602
64	conserjería	0,03644	0,323	79,102	7,018
65	tablón	0,03631	0,296	79,398	6,433
66	clasificador	0,03539	0,296	79,693	6,433
67	pasillo	0,03526	0,242	79,935	5,263
68	lapicero	0,03420	0,269	80,204	5,848
69	conserje	0,03366	0,269	80,473	5,848
70	hoja	0,03300	0,269	80,742	5,848
71	televisión	0,03041	0,269	81,011	5,848
72	director	0,02985	0,269	81,280	5,848
73	horario	0,02901	0,269	81,549	5,848
74	taquilla	0,02796	0,269	81,818	5,848
75	subrayador	0,02637	0,215	82,033	4,678
76	archivador	0,02605	0,269	82,302	5,848
77	plastilina	0,02555	0,269	82,571	5,848
78	internet	0,02515	0,242	82,813	5,263
79	salón de actos	0,02429	0,215	83,029	4,678
80	teclado	0,02253	0,161	83,190	3,509
81	corcho	0,02221	0,215	83,405	4,678
82	sala de profesores	0,02213	0,188	83,593	4,094
83	retroproyector	0,02210	0,215	83,809	4,678
84	deber	0,02061	0,215	84,024	4,678
85	luz	0,02017	0,188	84,212	4,094
86	póster	0,01987	0,188	84,400	4,094
87	estudiante	0,01986	0,134	84,535	2,924
88	radio	0,01961	0,188	84,723	4,094
89	comedor	0,01889	0,188	84,911	4,094
90	aseo	0,01851	0,134	85,046	2,924
91	jefe de estudios	0,01833	0,134	85,180	2,924
92	balón	0,01803	0,188	85,368	4,094
93	Post-it	0,01798	0,161	85,530	3,509

94	asignatura	0,01780	0,134	85,664	2,924
95	libro de texto	0,01777	0,108	85,772	2,339
96	secretaría	0,01769	0,161	85,933	3,509
97	cartuchera	0,01767	0,134	86,068	2,924
98	percha	0,01759	0,134	86,202	2,924
99	clip	0,01651	0,188	86,391	4,094
100	cartera	0,01622	0,161	86,552	3,509
101	ratón	0,01605	0,108	86,659	2,339
102	compañero	0,01576	0,108	86,767	2,339
103	pared	0,01527	0,108	86,875	2,339
104	colegio	0,01522	0,108	86,982	2,339
105	encerado	0,01498	0,108	87,090	2,339
106	altavoz	0,01488	0,134	87,224	2,924
107	escuela	0,01429	0,081	87,305	1,754
108	claustro	0,01428	0,081	87,386	1,754
109	techo	0,01423	0,108	87,493	2,339
110	pintura	0,01363	0,134	87,628	2,924
111	tablón de anuncios	0,01339	0,108	87,735	2,339
112	enciclopedia	0,01338	0,108	87,843	2,339
113	calendario	0,01330	0,134	87,977	2,924
114	atril	0,01329	0,108	88,085	2,339
115	material escolar	0,01323	0,081	88,166	1,754
116	servicio	0,01319	0,108	88,273	2,339
117	cuadernillo	0,01318	0,134	88,408	2,924
118	uniforme	0,01287	0,134	88,542	2,924
119	timbre	0,01284	0,134	88,677	2,924
120	persiana	0,01267	0,134	88,811	2,924
121	sala de ordenadores	0,01242	0,108	88,919	2,339
122	maestro	0,01242	0,081	88,999	1,754
123	corrector	0,01239	0,108	89,107	2,339
124	matemáticas	0,01232	0,134	89,242	2,924
125	mesa del profesor	0,01197	0,081	89,322	1,754
126	estudiar	0,01189	0,108	89,430	2,339
127	fotocopiadora	0,01158	0,134	89,564	2,924
128	laboratorio	0,01140	0,108	89,672	2,339
129	pista deportiva	0,01133	0,108	89,779	2,339
130	cajón	0,01107	0,108	89,887	2,339
131	apunte	0,01104	0,108	89,995	2,339
132	suelo	0,01073	0,081	90,075	1,754
133	papel de celofán	0,01069	0,081	90,156	1,754
134	chincheta	0,01064	0,134	90,290	2,924
135	estrado	0,01057	0,081	90,371	1,754
136	(rotulador)				
	fluorescente	0,01023	0,108	90,479	2,339
137	pincel	0,00981	0,108	90,586	2,339
138	cancha de				
	baloncesto	0,00966	0,081	90,667	1,754
139	juguete	0,00961	0,108	90,775	2,339
140	mina	0,00945	0,081	90,855	1,754

141	témpera	0,00942	0,108	90,963	2,339
142	copistería	0,00937	0,081	91,044	1,754
143	departamento	0,00910	0,081	91,124	1,754
144	centro TIC	0,00904	0,081	91,205	1,754
145	amigo	0,00886	0,081	91,286	1,754
146	fotocopia	0,00880	0,108	91,393	2,339
147	banqueta	0,00878	0,054	91,447	1,170
148	juego	0,00868	0,108	91,555	2,339
149	CD(-ROM)	0,00863	0,081	91,635	1,754
150	tinta	0,00834	0,108	91,743	2,339
151	dibujo	0,00833	0,081	91,824	1,754
152	sillón	0,00830	0,081	91,904	1,754
153	estante	0,00822	0,054	91,958	1,170
154	bocadillo	0,00820	0,054	92,012	1,170
155	acuarela	0,00805	0,081	92,093	1,754
156	física	0,00799	0,108	92,200	2,339
157	enchufe	0,00781	0,081	92,281	1,754
158	vídeo	0,00779	0,054	92,335	1,170
159	sala de informática	0,00770	0,081	92,415	1,754
160	instituto	0,00767	0,054	92,469	1,170
161	castigo	0,00764	0,081	92,550	1,754
162	colchoneta	0,00758	0,081	92,630	1,754
163	grapa	0,00732	0,054	92,684	1,170
164	universidad	0,00731	0,054	92,738	1,170
165	taburete	0,00728	0,054	92,792	1,170
166	taller	0,00703	0,054	92,846	1,170
167	nota	0,00689	0,081	92,926	1,754
168	globo terráqueo	0,00673	0,081	93,007	1,754
169	banca	0,00664	0,054	93,061	1,170
170	extintor	0,00658	0,054	93,115	1,170
171	pelota	0,00658	0,054	93,168	1,170
172	secretario	0,00633	0,054	93,222	1,170
173	lámina	0,00627	0,054	93,276	1,170
174	lengua	0,00621	0,081	93,357	1,754
175	gimnasia	0,00617	0,081	93,437	1,754
176	gafa	0,00607	0,081	93,518	1,754
177	autobús (escolar)	0,00585	0,054	93,572	1,170
178	pantalla digital	0,00585	0,027	93,599	0,585
179	diapositiva	0,00571	0,054	93,653	1,170
180	lavabo	0,00571	0,054	93,706	1,170
181	lista	0,00567	0,054	93,760	1,170
182	biología	0,00561	0,054	93,814	1,170
183	ejercicio	0,00549	0,054	93,868	1,170
184	atlas	0,00546	0,054	93,921	1,170
185	transparencia	0,00544	0,054	93,975	1,170
186	química	0,00538	0,081	94,056	1,754
187	reproductor	0,00532	0,054	94,110	1,170
188	bloc de notas	0,00531	0,027	94,137	0,585
189	pizarra blanca	0,00531	0,027	94,164	0,585

190	pizarra tradicional	0,00531	0,027	94,190	0,585
191	pupilo	0,00531	0,027	94,217	0,585
192	celo	0,00509	0,054	94,271	1,170
193	bombilla	0,00506	0,027	94,298	0,585
194	seminario	0,00506	0,027	94,325	0,585
195	urinario	0,00506	0,027	94,352	0,585
196	cafetería	0,00503	0,054	94,406	1,170
197	caja	0,00490	0,054	94,459	1,170
198	papel pinocho	0,00483	0,054	94,513	1,170
199	cajonera	0,00483	0,027	94,540	0,585
200	Pilot	0,00483	0,027	94,567	0,585
201	respaldo	0,00483	0,027	94,594	0,585
202	canasta	0,00478	0,054	94,648	1,170
203	actividad	0,00461	0,054	94,701	1,170
204	aire acondicionado	0,00460	0,027	94,728	0,585
205	funda	0,00460	0,027	94,755	0,585
206	imbornal	0,00460	0,027	94,782	0,585
207	jefatura de estudios	0,00460	0,027	94,809	0,585
208	lápiz de madera	0,00460	0,027	94,836	0,585
209	niño	0,00460	0,054	94,890	1,170
210	escalera	0,00457	0,054	94,944	1,170
211	control	0,00451	0,054	94,997	1,170
212	educación plástica y visual	0,00449	0,054	95,051	1,170
213	parte	0,00445	0,054	95,105	1,170
214	campana	0,00439	0,027	95,132	0,585
215	dintel	0,00439	0,027	95,159	0,585
216	Fixo	0,00439	0,027	95,186	0,585
217	iluminación	0,00439	0,027	95,212	0,585
218	historia	0,00432	0,054	95,266	1,170
219	materia	0,00418	0,027	95,293	0,585
220	página	0,00418	0,027	95,320	0,585
221	trimestre	0,00418	0,027	95,347	0,585
222	forro	0,00399	0,027	95,374	0,585
223	mural	0,00399	0,027	95,401	0,585
224	ordenador personal	0,00399	0,027	95,428	0,585
225	reproductor de vídeo	0,00399	0,027	95,455	0,585
226	administrativo	0,00380	0,027	95,481	0,585
227	boca de incendio equipada	0,00380	0,027	95,508	0,585
228	cama	0,00380	0,027	95,535	0,585
229	cortina	0,00380	0,027	95,562	0,585
230	institución	0,00380	0,027	95,589	0,585
231	porta-ángulos	0,00380	0,027	95,616	0,585
232	ropero	0,00380	0,027	95,643	0,585
233	televisor	0,00380	0,027	95,670	0,585
234	transportador	0,00380	0,027	95,697	0,585
235	aula de convivencia	0,00362	0,027	95,724	0,585

236	edificio	0,00362	0,027	95,750	0,585
237	esquema	0,00362	0,027	95,777	0,585
238	inglés	0,00362	0,027	95,804	0,585
239	marcador	0,00362	0,027	95,831	0,585
240	monja	0,00362	0,027	95,858	0,585
241	netbook	0,00362	0,027	95,885	0,585
242	suspenso	0,00362	0,027	95,912	0,585
243	aprobado	0,00345	0,027	95,939	0,585
244	consejo escolar	0,00345	0,027	95,966	0,585
245	fax	0,00345	0,027	95,992	0,585
246	francés	0,00345	0,027	96,019	0,585
247	mapamundi	0,00345	0,027	96,046	0,585
248	teléfono	0,00345	0,027	96,073	0,585
249	temario	0,00345	0,027	96,100	0,585
250	música	0,00343	0,054	96,154	1,170
251	AMPA	0,00329	0,027	96,181	0,585
252	ascensor	0,00329	0,027	96,208	0,585
253	baldosa	0,00329	0,027	96,235	0,585
254	centro bilingüe	0,00329	0,027	96,261	0,585
255	chándal	0,00329	0,027	96,288	0,585
256	dirección	0,00329	0,027	96,315	0,585
257	lengua castellana y literatura	0,00329	0,027	96,342	0,585
258	libro de lectura	0,00329	0,027	96,369	0,585
259	quitagrapas	0,00329	0,027	96,396	0,585
260	repaso	0,00329	0,027	96,423	0,585
261	concurso	0,00314	0,027	96,450	0,585
262	escuela pública	0,00314	0,027	96,477	0,585
263	fútbol	0,00314	0,027	96,503	0,585
264	panel	0,00314	0,027	96,530	0,585
265	rampa para minusválidos	0,00314	0,027	96,557	0,585
266	tenis	0,00314	0,027	96,584	0,585
267	bigotera	0,00299	0,027	96,611	0,585
268	bloc	0,00299	0,027	96,638	0,585
269	bola del mundo	0,00299	0,027	96,665	0,585
270	cuadro	0,00299	0,027	96,692	0,585
271	escuela concertada	0,00299	0,027	96,719	0,585
272	geología	0,00299	0,027	96,746	0,585
273	pendrive	0,00299	0,027	96,772	0,585
274	test	0,00299	0,027	96,799	0,585
275	transportador de ángulos	0,00299	0,027	96,826	0,585
276	transporte escolar	0,00299	0,027	96,853	0,585
277	despacho	0,00292	0,054	96,907	1,170
278	aprendizaje	0,00285	0,027	96,934	0,585
279	Biblia	0,00285	0,027	96,961	0,585

280	bolso	0,00285	0,027	96,988	0,585
281	escuela privada	0,00285	0,027	97,015	0,585
282	latín	0,00285	0,027	97,041	0,585
283	manual	0,00285	0,027	97,068	0,585
284	váter	0,00285	0,027	97,095	0,585
285	escalímetro	0,00272	0,027	97,122	0,585
286	griego	0,00272	0,027	97,149	0,585
287	jardín	0,00272	0,027	97,176	0,585
288	método	0,00272	0,027	97,203	0,585
289	paragüero	0,00272	0,027	97,230	0,585
290	profesorado	0,00272	0,027	97,257	0,585
291	reloj	0,00272	0,027	97,283	0,585
292	revista	0,00272	0,027	97,310	0,585
293	semicírculo				
	graduado	0,00272	0,027	97,337	0,585
294	tablón de corcho	0,00272	0,027	97,364	0,585
295	alumnado	0,00259	0,027	97,391	0,585
296	barro	0,00259	0,027	97,418	0,585
297	cantina	0,00259	0,027	97,445	0,585
298	capilla	0,00259	0,027	97,472	0,585
299	dossier	0,00259	0,027	97,499	0,585
300	espejo	0,00259	0,027	97,526	0,585
301	excursión	0,00259	0,027	97,552	0,585
302	filosofía	0,00259	0,027	97,579	0,585
303	geografía	0,00259	0,027	97,606	0,585
304	lenguaje	0,00259	0,027	97,633	0,585
305	manualidad	0,00259	0,027	97,660	0,585
306	unidad didáctica	0,00259	0,027	97,687	0,585
307	ciencias naturales	0,00247	0,027	97,714	0,585
308	curso académico	0,00247	0,027	97,741	0,585
309	enemigo	0,00247	0,027	97,768	0,585
310	hall	0,00247	0,027	97,795	0,585
311	papel de seda	0,00247	0,027	97,821	0,585
312	plumier	0,00247	0,027	97,848	0,585
313	polideportivo	0,00247	0,027	97,875	0,585
314	potro	0,00247	0,027	97,902	0,585
315	producto químico	0,00247	0,027	97,929	0,585
316	programación	0,00247	0,027	97,956	0,585
317	suspender	0,00247	0,027	97,983	0,585
318	tecnología	0,00247	0,027	98,010	0,585
319	aprobar	0,00235	0,027	98,037	0,585
320	columpio	0,00235	0,027	98,063	0,585
321	impresora	0,00235	0,027	98,090	0,585
322	mesa de dibujo	0,00235	0,027	98,117	0,585
323	pegatina	0,00235	0,027	98,144	0,585
324	personal de limpieza	0,00235	0,027	98,171	0,585

325	religión	0,00235	0,027	98,198	0,585
326	sala de música	0,00235	0,027	98,225	0,585
327	abrigo	0,00224	0,027	98,252	0,585
328	cura	0,00224	0,027	98,279	0,585
329	educación física	0,00224	0,027	98,306	0,585
330	estudio	0,00224	0,027	98,332	0,585
331	librería	0,00224	0,027	98,359	0,585
332	matrícula	0,00224	0,027	98,386	0,585
333	orientador	0,00224	0,027	98,413	0,585
334	tubo fluorescente	0,00224	0,027	98,440	0,585
335	alfombra	0,00214	0,027	98,467	0,585
336	campus	0,00214	0,027	98,494	0,585
337	comportamiento	0,00214	0,027	98,521	0,585
338	cuadrícula	0,00214	0,027	98,548	0,585
339	enseñar	0,00214	0,027	98,575	0,585
340	guardería	0,00214	0,027	98,601	0,585
341	líquido corrector	0,00214	0,027	98,628	0,585
342	portero	0,00214	0,027	98,655	0,585
343	afilador	0,00204	0,027	98,682	0,585
344	aprender	0,00204	0,027	98,709	0,585
345	cartilla	0,00204	0,027	98,736	0,585
346	diversión	0,00204	0,027	98,763	0,585
347	escáner	0,00204	0,027	98,790	0,585
348	imagen	0,00204	0,027	98,817	0,585
349	lectura	0,00204	0,027	98,843	0,585
350	portafolio	0,00204	0,027	98,870	0,585
351	probeta	0,00204	0,027	98,897	0,585
352	reunión	0,00204	0,027	98,924	0,585
353	trabajo	0,00204	0,027	98,951	0,585
354	agua	0,00194	0,027	98,978	0,585
355	caligrafía	0,00194	0,027	99,005	0,585
356	grande	0,00194	0,027	99,032	0,585
357	Séneca	0,00194	0,027	99,059	0,585
358	tarea	0,00194	0,027	99,086	0,585
359	auricular	0,00185	0,027	99,112	0,585
360	fuelle	0,00185	0,027	99,139	0,585
361	parte de faltas	0,00185	0,027	99,166	0,585
362	tutor	0,00185	0,027	99,193	0,585
363	butaca	0,00177	0,027	99,220	0,585
364	cable	0,00177	0,027	99,247	0,585
365	libro del profesor	0,00177	0,027	99,274	0,585
366	sanción	0,00177	0,027	99,301	0,585
367	separador	0,00177	0,027	99,328	0,585
368	bloc de dibujo	0,00168	0,027	99,354	0,585
369	celofán	0,00160	0,027	99,381	0,585
370	conocimiento	0,00160	0,027	99,408	0,585

371	cuerda	0,00160	0,027	99,435	0,585
372	resumen	0,00160	0,027	99,462	0,585
373	diario	0,00153	0,027	99,489	0,585
374	ética	0,00153	0,027	99,516	0,585
375	papel milimetrado	0,00153	0,027	99,543	0,585
376	película	0,00153	0,027	99,570	0,585
377	pica	0,00153	0,027	99,597	0,585
378	raqueta	0,00153	0,027	99,623	0,585
379	redacción	0,00153	0,027	99,650	0,585
380	muñeco	0,00146	0,027	99,677	0,585
381	papel vegetal	0,00146	0,027	99,704	0,585
382	portería	0,00146	0,027	99,731	0,585
383	aro	0,00139	0,027	99,758	0,585
384	babero	0,00139	0,027	99,785	0,585
385	papel cebolla	0,00139	0,027	99,812	0,585
386	grupo	0,00132	0,027	99,839	0,585
387	reproductor de DVD	0,00132	0,027	99,866	0,585
388	tambor	0,00126	0,027	99,892	0,585
389	desayuno	0,00120	0,027	99,919	0,585
390	madrugar	0,00120	0,027	99,946	0,585
391	pandereta	0,00115	0,027	99,973	0,585
392	guitarra	0,00109	0,027	100,000	0,585

1.2.5. La ciudad

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	calle	0,54776	2,940	2,940	63,743
2	coche	0,50193	2,913	5,852	63,158
3	edificio	0,49611	3,020	8,873	65,497
4	semáforo	0,40953	2,616	11,489	56,725
5	parque	0,38495	2,670	14,159	57,895
6	tienda	0,38147	2,751	16,909	59,649
7	carretera	0,29351	1,807	18,716	39,181
8	casa	0,27729	1,807	20,523	39,181
9	plaza	0,26690	1,807	22,330	39,181
10	centro comercial	0,23725	1,780	24,110	38,596
11	(auto)bus	0,23166	1,645	25,755	35,673
12	restaurante	0,19179	1,510	27,265	32,749
13	acera	0,19128	1,375	28,641	29,825
14	bar	0,18353	1,456	30,097	31,579
15	avenida	0,18344	1,133	31,230	24,561
16	museo	0,18254	1,268	32,497	27,485
17	ayuntamiento	0,18215	1,268	33,765	27,485
18	cine	0,17308	1,321	35,086	28,655
19	hospital	0,15957	1,214	36,300	26,316
20	gente	0,15670	0,998	37,298	21,637
21	tráfico	0,15439	0,998	38,296	21,637
22	banco	0,15364	1,241	39,536	26,901
23	teatro	0,14916	1,106	40,642	23,977
24	árbol	0,14364	1,052	41,694	22,807
25	iglesia	0,14265	1,052	42,745	22,807
26	rotonda	0,13413	0,998	43,743	21,637
27	colegio	0,13254	1,025	44,768	22,222
28	universidad	0,13232	1,106	45,874	23,977
29	metro	0,13208	0,971	46,845	21,053
30	supermercado	0,11842	0,944	47,789	20,468
31	jardín	0,11733	0,890	48,679	19,298
32	monumento	0,11584	0,836	49,515	18,129
33	aeropuerto	0,11264	0,890	50,405	19,298
34	paso de peatones	0,11108	0,809	51,214	17,544
35	fuelle	0,10379	0,836	52,050	18,129
36	farola	0,10210	0,728	52,778	15,789
37	piso	0,09707	0,701	53,479	15,205
38	panadería	0,09704	0,782	54,261	16,959
39	taxi	0,09244	0,701	54,962	15,205
40	escuela	0,09205	0,728	55,690	15,789
41	puente	0,09148	0,728	56,419	15,789
42	catedral	0,08922	0,647	57,066	14,035
43	ruido	0,08896	0,620	57,686	13,450
44	moto(cicleta)	0,08895	0,566	58,252	12,281
45	peatón	0,07855	0,512	58,765	11,111

46	estación de tren	0,07837	0,566	59,331	12,281
47	estación	0,07565	0,593	59,924	12,865
48	puerto	0,07529	0,566	60,491	12,281
49	bici(cleta)	0,07485	0,512	61,003	11,111
50	playa	0,07422	0,620	61,624	13,450
51	rascacielos	0,07251	0,566	62,190	12,281
52	barrio	0,07248	0,566	62,756	12,281
53	contaminación	0,07056	0,512	63,269	11,111
54	cafetería	0,07024	0,539	63,808	11,696
55	discoteca	0,06725	0,593	64,401	12,865
56	instituto	0,06665	0,566	64,968	12,281
57	paseo marítimo	0,06372	0,485	65,453	10,526
58	policía	0,06311	0,485	65,939	10,526
59	aparcamiento	0,05957	0,512	66,451	11,111
60	tren	0,05925	0,458	66,909	9,942
61	papelera	0,05903	0,512	67,422	11,111
62	centro	0,05882	0,378	67,799	8,187
63	comercio	0,05832	0,458	68,258	9,942
64	comisaría (de				
	policía)	0,05802	0,485	68,743	10,526
65	pub	0,05591	0,485	69,229	10,526
66	oficina	0,05358	0,405	69,633	8,772
67	persona	0,05300	0,351	69,984	7,602
68	mercado	0,04939	0,378	70,361	8,187
69	carnicería	0,04673	0,405	70,766	8,772
70	biblioteca	0,04662	0,405	71,170	8,772
71	parada de autobús	0,04642	0,378	71,548	8,187
72	librería	0,04578	0,378	71,926	8,187
73	atasco	0,04223	0,297	72,222	6,433
74	señal	0,03946	0,297	72,519	6,433
75	luz	0,03898	0,297	72,816	6,433
76	farmacia	0,03832	0,297	73,112	6,433
77	estatua	0,03672	0,270	73,382	5,848
78	polución	0,03659	0,270	73,652	5,848
79	paseo	0,03623	0,270	73,921	5,848
80	señal de tráfico	0,03606	0,324	74,245	7,018
81	Correos	0,03585	0,243	74,488	5,263
82	río	0,03525	0,297	74,784	6,433
83	perro	0,03524	0,324	75,108	7,018
84	pescadería	0,03518	0,297	75,405	6,433
85	bulevar	0,03411	0,216	75,620	4,678
86	guardería	0,03365	0,324	75,944	7,018
87	quiosco	0,03361	0,270	76,214	5,848
88	callejón	0,03250	0,216	76,429	4,678
89	apartamento	0,03186	0,243	76,672	5,263
90	autovía	0,03180	0,216	76,888	4,678
91	polideportivo	0,03144	0,270	77,157	5,848
92	estrés	0,03086	0,189	77,346	4,094
93	estación de autobús	0,02877	0,216	77,562	4,678

94	camión	0,02834	0,189	77,751	4,094
95	palacio	0,02718	0,216	77,967	4,678
96	obra	0,02717	0,243	78,209	5,263
97	bloque de pisos	0,02715	0,189	78,398	4,094
98	paso de cebra	0,02659	0,189	78,587	4,094
99	parking	0,02655	0,243	78,830	5,263
100	tranvía	0,02653	0,243	79,072	5,263
101	carril bici	0,02609	0,189	79,261	4,094
102	trabajo	0,02546	0,216	79,477	4,678
103	urbanización	0,02476	0,243	79,720	5,263
104	hotel	0,02373	0,216	79,935	4,678
105	basura	0,02252	0,189	80,124	4,094
106	bloque	0,02173	0,135	80,259	2,924
107	columpio	0,02166	0,189	80,448	4,094
108	cruce	0,02153	0,162	80,609	3,509
109	autopista	0,02141	0,135	80,744	2,924
110	centro histórico	0,02132	0,135	80,879	2,924
111	mar	0,02112	0,189	81,068	4,094
112	mercadillo	0,02087	0,189	81,257	4,094
113	cajero (automático)	0,02041	0,162	81,419	3,509
114	piscina	0,01931	0,189	81,607	4,094
115	alcazaba	0,01902	0,135	81,742	2,924
116	niño	0,01895	0,189	81,931	4,094
117	oficina del INEM	0,01846	0,108	82,039	2,339
118	peluquería	0,01841	0,162	82,201	3,509
119	centro de salud	0,01831	0,189	82,389	4,094
120	avión	0,01827	0,108	82,497	2,339
121	galería de arte	0,01827	0,108	82,605	2,339
122	casco histórico	0,01794	0,108	82,713	2,339
123	escaparate	0,01766	0,135	82,848	2,924
124	estadio	0,01718	0,189	83,037	4,094
125	gasolinera	0,01664	0,135	83,172	2,924
126	humo	0,01641	0,108	83,279	2,339
127	gimnasio	0,01613	0,135	83,414	2,924
128	casco antiguo	0,01576	0,081	83,495	1,754
129	campo de fútbol	0,01531	0,135	83,630	2,924
130	gato	0,01527	0,135	83,765	2,924
131	floristería	0,01490	0,135	83,900	2,924
132	chalé	0,01482	0,162	84,061	3,509
133	asfalto	0,01464	0,108	84,169	2,339
134	contenedor de basura	0,01452	0,108	84,277	2,339
135	tetería	0,01440	0,135	84,412	2,924
136	túnel	0,01433	0,135	84,547	2,924
137	ferretería	0,01403	0,135	84,682	2,924
138	aglomeración	0,01381	0,108	84,790	2,339
139	chiringuito	0,01381	0,135	84,924	2,924
140	transporte	0,01376	0,081	85,005	1,754
141	calzada	0,01370	0,135	85,140	2,924

142	facultad	0,01364	0,135	85,275	2,924
143	caravana	0,01328	0,108	85,383	2,339
144	parada	0,01307	0,108	85,491	2,339
145	pueblo	0,01304	0,081	85,572	1,754
146	frutería	0,01303	0,135	85,707	2,924
147	montaña	0,01281	0,108	85,814	2,339
148	zapatería	0,01280	0,162	85,976	3,509
149	zoológico	0,01277	0,108	86,084	2,339
150	prisa	0,01267	0,081	86,165	1,754
151	transporte público	0,01259	0,108	86,273	2,339
152	cementerio	0,01251	0,108	86,381	2,339
153	academia	0,01229	0,108	86,489	2,339
154	bufete de abogados	0,01221	0,081	86,570	1,754
155	zona residencial	0,01217	0,081	86,650	1,754
156	bombero	0,01215	0,108	86,758	2,339
157	ascensor	0,01207	0,081	86,839	1,754
158	planta	0,01207	0,135	86,974	2,924
159	transeúnte	0,01177	0,081	87,055	1,754
160	copistería	0,01177	0,108	87,163	2,339
161	pastelería	0,01172	0,108	87,271	2,339
162	confitería	0,01168	0,108	87,379	2,339
163	almacén	0,01164	0,081	87,460	1,754
164	lago	0,01162	0,081	87,540	1,754
165	grandes almacenes	0,01151	0,081	87,621	1,754
166	adosado	0,01146	0,108	87,729	2,339
167	turista	0,01145	0,108	87,837	2,339
168	manzana	0,01133	0,081	87,918	1,754
169	alcalde	0,01116	0,054	87,972	1,170
170	oficina de empleo	0,01115	0,054	88,026	1,170
171	fiesta	0,01114	0,081	88,107	1,754
172	cabina	0,01100	0,081	88,188	1,754
173	carril	0,01092	0,081	88,269	1,754
174	bullicio	0,01083	0,081	88,350	1,754
175	ópera	0,01056	0,081	88,430	1,754
176	jefatura de policía	0,01022	0,108	88,538	2,339
177	Escuela Oficial de Idiomas	0,01017	0,081	88,619	1,754
178	bordillo	0,01015	0,081	88,700	1,754
179	zona verde	0,01014	0,054	88,754	1,170
180	centro de transporte	0,01013	0,054	88,808	1,170
181	vivienda	0,01005	0,081	88,889	1,754
182	escalera automática	0,01003	0,081	88,970	1,754
183	alcantarilla	0,00996	0,081	89,051	1,754
184	obelisco	0,00995	0,081	89,132	1,754
185	polígono	0,00991	0,081	89,213	1,754
186	parque de bomberos	0,00966	0,081	89,293	1,754
187	charcutería	0,00956	0,054	89,347	1,170
188	cárcel	0,00955	0,108	89,455	2,339

189	ajetreo	0,00950	0,081	89,536	1,754
190	hostal	0,00938	0,108	89,644	2,339
191	suburbio	0,00937	0,054	89,698	1,170
192	ambulancia	0,00933	0,081	89,779	1,754
193	papelería	0,00932	0,081	89,860	1,754
194	plaza de toros	0,00929	0,081	89,941	1,754
195	contenedor de reciclaje	0,00908	0,081	90,022	1,754
196	vehículo	0,00905	0,054	90,076	1,170
197	castillo	0,00903	0,081	90,156	1,754
198	tienda de ropa	0,00902	0,108	90,264	2,339
199	suciedad	0,00894	0,081	90,345	1,754
200	Delegación de Hacienda	0,00887	0,081	90,426	1,754
201	callejuela	0,00886	0,054	90,480	1,170
202	galería	0,00886	0,054	90,534	1,170
203	cabina telefónica	0,00878	0,054	90,588	1,170
204	diversión	0,00877	0,054	90,642	1,170
205	capilla	0,00859	0,054	90,696	1,170
206	casa mata	0,00837	0,081	90,777	1,754
207	arcén	0,00830	0,054	90,831	1,170
208	clínica	0,00812	0,054	90,885	1,170
209	mercería	0,00806	0,081	90,965	1,754
210	monolito	0,00804	0,054	91,019	1,170
211	cartel	0,00803	0,054	91,073	1,170
212	negocio	0,00801	0,054	91,127	1,170
213	consultorio médico	0,00798	0,054	91,181	1,170
214	ciudadano	0,00792	0,054	91,235	1,170
215	accidente	0,00790	0,081	91,316	1,754
216	palacio de congresos	0,00779	0,054	91,370	1,170
217	pasaje	0,00778	0,054	91,424	1,170
218	campus (universitario)	0,00778	0,054	91,478	1,170
219	flor	0,00768	0,054	91,532	1,170
220	institución	0,00745	0,054	91,586	1,170
221	teatro romano	0,00717	0,054	91,640	1,170
222	pájaro	0,00717	0,081	91,721	1,754
223	calle peatonal	0,00711	0,081	91,802	1,754
224	agencia de viajes	0,00709	0,054	91,855	1,170
225	estación de metro	0,00698	0,054	91,909	1,170
226	habitante	0,00690	0,054	91,963	1,170
227	RENFE	0,00684	0,054	92,017	1,170
228	terrazza	0,00677	0,054	92,071	1,170
229	anuncio	0,00670	0,054	92,125	1,170
230	viandante	0,00665	0,054	92,179	1,170
231	juguetería	0,00664	0,054	92,233	1,170
232	juzgado	0,00663	0,054	92,287	1,170
233	ático	0,00659	0,054	92,341	1,170

234	tapería	0,00654	0,054	92,395	1,170
235	AVE	0,00652	0,054	92,449	1,170
236	circulación	0,00642	0,054	92,503	1,170
237	carril bus	0,00630	0,081	92,584	1,754
238	relojería	0,00628	0,054	92,638	1,170
239	línea de autobús	0,00613	0,054	92,691	1,170
240	estanco	0,00609	0,054	92,745	1,170
241	McDonald's	0,00604	0,054	92,799	1,170
242	torre	0,00604	0,054	92,853	1,170
243	marquesina	0,00604	0,054	92,907	1,170
244	compra	0,00601	0,054	92,961	1,170
245	heladería	0,00595	0,081	93,042	1,754
246	mascota	0,00585	0,054	93,096	1,170
247	consulado	0,00584	0,054	93,150	1,170
248	embajada	0,00584	0,054	93,204	1,170
249	anciano	0,00576	0,054	93,258	1,170
250	cartel publicitario	0,00576	0,054	93,312	1,170
251	fábrica	0,00569	0,081	93,393	1,754
252	individuo	0,00557	0,027	93,420	0,585
253	verja	0,00557	0,027	93,447	0,585
254	glorieta	0,00551	0,081	93,528	1,754
255	zona recreativa	0,00549	0,054	93,581	1,170
256	centro cívico	0,00531	0,027	93,608	0,585
257	construcción	0,00531	0,027	93,635	0,585
258	esquina	0,00531	0,027	93,662	0,585
259	casa adosada	0,00531	0,054	93,716	1,170
260	albergue	0,00519	0,054	93,770	1,170
261	acequia	0,00506	0,027	93,797	0,585
262	basílica	0,00506	0,027	93,824	0,585
263	dúplex	0,00506	0,027	93,851	0,585
264	muelle	0,00506	0,027	93,878	0,585
265	problema	0,00506	0,027	93,905	0,585
266	señalización	0,00506	0,027	93,932	0,585
267	tierra	0,00506	0,027	93,959	0,585
268	aduana	0,00483	0,027	93,986	0,585
269	Alliance Française	0,00483	0,027	94,013	0,585
270	boca de metro	0,00483	0,027	94,040	0,585
271	buzón	0,00483	0,027	94,067	0,585
272	callejón sin salida	0,00483	0,027	94,094	0,585
273	cuadro	0,00483	0,027	94,121	0,585
274	escalinata	0,00483	0,027	94,148	0,585
275	plano	0,00483	0,027	94,175	0,585
276	rambla	0,00483	0,027	94,202	0,585
277	vía	0,00483	0,027	94,229	0,585
278	suelo	0,00475	0,054	94,283	1,170
279	tobogán	0,00471	0,054	94,337	1,170

280	alameda	0,00460	0,027	94,364	0,585
281	barco	0,00460	0,027	94,391	0,585
282	claxon	0,00460	0,027	94,417	0,585
283	pabellón	0,00460	0,027	94,444	0,585
284	pelea	0,00460	0,027	94,471	0,585
285	sirena	0,00460	0,027	94,498	0,585
286	zona azul	0,00460	0,027	94,525	0,585
287	ciclista	0,00439	0,027	94,552	0,585
288	clínica veterinaria	0,00439	0,027	94,579	0,585
289	conservatorio	0,00439	0,027	94,606	0,585
290	hipermercado	0,00439	0,027	94,633	0,585
291	servicio	0,00439	0,027	94,660	0,585
292	vecindario	0,00439	0,027	94,687	0,585
293	vecino	0,00439	0,027	94,714	0,585
294	zona peatonal	0,00439	0,027	94,741	0,585
295	gran superficie	0,00427	0,054	94,795	1,170
296	cajero (automático)				
	señal	0,00418	0,027	94,822	0,585
297	circunvalación	0,00418	0,027	94,849	0,585
298	club nocturno	0,00418	0,027	94,876	0,585
299	EMT	0,00418	0,027	94,903	0,585
300	energía	0,00418	0,027	94,930	0,585
301	grúa	0,00418	0,027	94,957	0,585
302	perrera	0,00418	0,027	94,984	0,585
303	rebajas	0,00418	0,027	95,011	0,585
304	turismo	0,00418	0,027	95,038	0,585
305	tienda de ultramarinos	0,00409	0,054	95,092	1,170
306	baldosa	0,00407	0,054	95,146	1,170
307	agencia	0,00399	0,027	95,173	0,585
308	cibercafé	0,00399	0,027	95,200	0,585
309	cielo	0,00399	0,027	95,227	0,585
310	excavadora	0,00399	0,027	95,254	0,585
311	foro romano	0,00399	0,027	95,280	0,585
312	gerencia	0,00399	0,027	95,307	0,585
313	peaje	0,00399	0,027	95,334	0,585
314	película	0,00399	0,027	95,361	0,585
315	periferia	0,00399	0,027	95,388	0,585
316	stop	0,00399	0,027	95,415	0,585
317	veterinario	0,00399	0,027	95,442	0,585
318	óptica	0,00385	0,054	95,496	1,170
319	mirador	0,00380	0,054	95,550	1,170
320	bazar	0,00380	0,027	95,577	0,585
321	ceda el paso	0,00380	0,027	95,604	0,585
322	centro de estética	0,00380	0,027	95,631	0,585
323	concierto	0,00380	0,027	95,658	0,585

324	cornisa	0,00380	0,027	95,685	0,585
325	establecimiento	0,00380	0,027	95,712	0,585
326	estadio de fútbol	0,00380	0,027	95,739	0,585
327	gestoría	0,00380	0,027	95,766	0,585
328	grito	0,00380	0,027	95,793	0,585
329	moda	0,00380	0,027	95,820	0,585
330	outlet	0,00380	0,027	95,847	0,585
331	redonda	0,00380	0,027	95,874	0,585
332	universidad pública	0,00380	0,027	95,901	0,585
333	ansiedad	0,00362	0,027	95,928	0,585
334	asiento	0,00362	0,027	95,955	0,585
335	Burger King	0,00362	0,027	95,982	0,585
336	clínica privada	0,00362	0,027	96,009	0,585
337	coche de caballos	0,00362	0,027	96,036	0,585
338	contenedor	0,00362	0,027	96,063	0,585
339	entretenimiento	0,00362	0,027	96,090	0,585
340	helipuerto	0,00362	0,027	96,117	0,585
341	notaría	0,00362	0,027	96,143	0,585
342	patín	0,00362	0,027	96,170	0,585
343	radar	0,00362	0,027	96,197	0,585
344	universidad privada	0,00362	0,027	96,224	0,585
345	boca de incendio	0,00345	0,027	96,251	0,585
346	camión de la basura	0,00345	0,027	96,278	0,585
347	centro concertado	0,00345	0,027	96,305	0,585
348	escalera	0,00345	0,027	96,332	0,585
349	inmueble	0,00345	0,027	96,359	0,585
350	raqueta	0,00345	0,027	96,386	0,585
351	bolera	0,00329	0,027	96,413	0,585
352	campo	0,00329	0,027	96,440	0,585
353	centro social	0,00329	0,027	96,467	0,585
354	ciudad	0,00329	0,027	96,494	0,585
355	cuartel de la Guardia Civil	0,00329	0,027	96,521	0,585
356	droguería	0,00329	0,027	96,548	0,585
357	Madrid	0,00329	0,027	96,575	0,585
358	ministerio	0,00329	0,027	96,602	0,585
359	organismo público	0,00329	0,027	96,629	0,585
360	patio	0,00329	0,027	96,656	0,585
361	pito	0,00329	0,027	96,683	0,585
362	antena	0,00329	0,054	96,737	1,170
363	alumbrado	0,00314	0,027	96,764	0,585
364	andén	0,00314	0,027	96,791	0,585
365	bandera	0,00314	0,027	96,818	0,585
366	Barcelona	0,00314	0,027	96,845	0,585
367	centro diurno	0,00314	0,027	96,872	0,585
368	decomisos	0,00314	0,027	96,899	0,585

369	dirección	0,00314	0,027	96,926	0,585
370	garaje	0,00314	0,027	96,953	0,585
371	joyería	0,00314	0,027	96,980	0,585
372	medio de transporte	0,00314	0,027	97,006	0,585
373	perfumería	0,00314	0,027	97,033	0,585
374	sonido	0,00314	0,027	97,060	0,585
375	ventana	0,00314	0,027	97,087	0,585
376	acueducto	0,00299	0,027	97,114	0,585
377	casa hermandad	0,00299	0,027	97,141	0,585
378	conductor	0,00299	0,027	97,168	0,585
379	delegación ministerial	0,00299	0,027	97,195	0,585
380	desagüe	0,00299	0,027	97,222	0,585
381	estanque	0,00299	0,027	97,249	0,585
382	estudio fotográfico	0,00299	0,027	97,276	0,585
383	faro	0,00299	0,027	97,303	0,585
384	inmobiliaria	0,00299	0,027	97,330	0,585
385	instituto de enseñanza secundaria	0,00299	0,027	97,357	0,585
386	Málaga	0,00299	0,027	97,384	0,585
387	pizzería	0,00299	0,027	97,411	0,585
388	ronda	0,00299	0,027	97,438	0,585
389	alcantarillado	0,00285	0,027	97,465	0,585
390	badén	0,00285	0,027	97,492	0,585
391	camino	0,00285	0,027	97,519	0,585
392	empresa	0,00285	0,027	97,546	0,585
393	golondrina	0,00285	0,027	97,573	0,585
394	hogar de jubilados	0,00285	0,027	97,600	0,585
395	lavandería	0,00285	0,027	97,627	0,585
396	nave industrial	0,00285	0,027	97,654	0,585
397	piscina municipal	0,00285	0,027	97,681	0,585
398	pozo	0,00285	0,027	97,708	0,585
399	puesto	0,00285	0,027	97,735	0,585
400	sede	0,00285	0,027	97,762	0,585
401	viejo	0,00285	0,027	97,789	0,585
402	abuelo	0,00272	0,027	97,816	0,585
403	capital	0,00272	0,027	97,843	0,585
404	concejal	0,00272	0,027	97,869	0,585
405	estudio	0,00272	0,027	97,896	0,585
406	gaviota	0,00272	0,027	97,923	0,585
407	jardín de infancia	0,00272	0,027	97,950	0,585
408	joven	0,00272	0,027	97,977	0,585
409	prisión	0,00272	0,027	98,004	0,585
410	valle	0,00272	0,027	98,031	0,585
411	barbería	0,00259	0,027	98,058	0,585

412	centro de belleza	0,00259	0,027	98,085	0,585
413	comunidad	0,00259	0,027	98,112	0,585
414	electricidad	0,00259	0,027	98,139	0,585
415	festival	0,00259	0,027	98,166	0,585
416	Junta de Distrito	0,00259	0,027	98,193	0,585
417	residencia de ancianos	0,00259	0,027	98,220	0,585
418	tiempo	0,00259	0,027	98,247	0,585
419	zona vip	0,00259	0,027	98,274	0,585
420	autobús urbano	0,00247	0,027	98,301	0,585
421	bar de tapas	0,00247	0,027	98,328	0,585
422	cultivo	0,00247	0,027	98,355	0,585
423	ente financiera	0,00247	0,027	98,382	0,585
424	guardia	0,00247	0,027	98,409	0,585
425	ocio	0,00247	0,027	98,436	0,585
426	palacete	0,00247	0,027	98,463	0,585
427	tintorería	0,00247	0,027	98,490	0,585
428	cervecería	0,00235	0,027	98,517	0,585
429	contador de agua	0,00235	0,027	98,544	0,585
430	cultura	0,00235	0,027	98,571	0,585
431	isleta	0,00235	0,027	98,598	0,585
432	jardinera	0,00235	0,027	98,625	0,585
433	nube	0,00235	0,027	98,652	0,585
434	Bolsa	0,00224	0,027	98,679	0,585
435	club	0,00224	0,027	98,706	0,585
436	El Retiro	0,00224	0,027	98,732	0,585
437	mediana	0,00224	0,027	98,759	0,585
438	seto	0,00224	0,027	98,786	0,585
439	afueras	0,00214	0,027	98,813	0,585
440	arbusto	0,00214	0,027	98,840	0,585
441	césped	0,00214	0,027	98,867	0,585
442	franquicia	0,00214	0,027	98,894	0,585
443	hamaca	0,00214	0,027	98,921	0,585
444	resbaladera	0,00214	0,027	98,948	0,585
445	salir	0,00214	0,027	98,975	0,585
446	videoclub	0,00214	0,027	99,002	0,585
447	arriate	0,00204	0,027	99,029	0,585
448	barrendero	0,00204	0,027	99,056	0,585
449	Cercanías	0,00204	0,027	99,083	0,585
450	cosa	0,00204	0,027	99,110	0,585
451	lámpara	0,00204	0,027	99,137	0,585
452	Plaza Mayor	0,00204	0,027	99,164	0,585
453	tanatorio	0,00204	0,027	99,191	0,585
454	ambulatorio	0,00194	0,027	99,218	0,585
455	Emasa	0,00194	0,027	99,245	0,585
456	hogar	0,00194	0,027	99,272	0,585

457	puesto de flores	0,00194	0,027	99,299	0,585
458	resto arqueológico	0,00194	0,027	99,326	0,585
459	sombra	0,00194	0,027	99,353	0,585
460	Endesa	0,00185	0,027	99,380	0,585
461	taberna	0,00185	0,027	99,407	0,585
462	tejado	0,00185	0,027	99,434	0,585
463	animal	0,00177	0,027	99,461	0,585
464	chimenea	0,00177	0,027	99,488	0,585
465	hogar de ancianos	0,00177	0,027	99,515	0,585
466	parada de taxi	0,00177	0,027	99,542	0,585
467	tensión	0,00177	0,027	99,569	0,585
468	boutique	0,00168	0,027	99,595	0,585
469	mendigo	0,00168	0,027	99,622	0,585
470	pasarela	0,00168	0,027	99,649	0,585
471	souvenir	0,00168	0,027	99,676	0,585
472	cable	0,00153	0,027	99,703	0,585
473	chabola	0,00153	0,027	99,730	0,585
474	Nueva York	0,00153	0,027	99,757	0,585
475	dinero	0,00146	0,027	99,784	0,585
476	muchedumbre	0,00146	0,027	99,811	0,585
477	cansancio	0,00139	0,027	99,838	0,585
478	psiquiátrico	0,00139	0,027	99,865	0,585
479	balcón	0,00132	0,027	99,892	0,585
480	orfanato	0,00132	0,027	99,919	0,585
481	cuartel	0,00126	0,027	99,946	0,585
482	amplitud	0,00120	0,027	99,973	0,585
483	Hiperronda	0,00115	0,027	100,000	0,585

1.2.6. Medios de transporte

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	coche	0,85702	5,983	5,983	92,398
2	avión	0,72650	5,907	11,889	91,228
3	tren	0,70709	5,680	17,569	87,719
4	bici(cleta)	0,66119	5,377	22,946	83,041
5	(auto)bus	0,65868	4,960	27,906	76,608
6	moto(cicleta)	0,59917	4,922	32,828	76,023
7	barco	0,59187	5,150	37,978	79,532
8	metro	0,47767	4,241	42,219	65,497
9	taxi	0,33026	2,764	44,983	42,690
10	tranvía	0,32729	2,953	47,936	45,614
11	helicóptero	0,30242	2,991	50,928	46,199
12	patín	0,27118	2,688	53,616	41,520
13	camión	0,27092	2,537	56,153	39,181
14	caballo	0,23809	2,461	58,614	38,012
15	furgoneta	0,19531	1,855	60,470	28,655
16	yate	0,19211	1,893	62,363	29,240
17	monopatín	0,18743	1,855	64,218	28,655
18	avioneta	0,14597	1,515	65,733	23,392
19	patinete	0,12943	1,363	67,096	21,053
20	crucero	0,11806	1,287	68,383	19,883
21	barca	0,10556	1,136	69,519	17,544
22	submarino	0,10256	1,060	70,579	16,374
23	burro	0,09935	1,022	71,602	15,789
24	ciclomotor	0,09063	0,795	72,397	12,281
25	coche de caballos	0,08806	0,947	73,343	14,620
26	globo	0,08723	0,909	74,252	14,035
27	lancha	0,07419	0,757	75,009	11,696
28	a pie	0,07413	0,757	75,767	11,696
29	velero	0,07116	0,757	76,524	11,696
30	autocar	0,06937	0,530	77,054	8,187
31	patera	0,06873	0,757	77,811	11,696
32	aeroplano	0,06855	0,644	78,455	9,942
33	ferry	0,06557	0,682	79,137	10,526
34	tráiler	0,06407	0,568	79,705	8,772
35	triciclo	0,06225	0,719	80,424	11,111
36	transatlántico	0,05990	0,682	81,106	10,526
37	AVE	0,05640	0,530	81,636	8,187
38	caravana	0,05434	0,530	82,166	8,187
39	camello	0,05136	0,606	82,772	9,357
40	tractor	0,04388	0,530	83,302	8,187
41	quad	0,04282	0,492	83,794	7,602
42	sidecar	0,04271	0,492	84,286	7,602
43	canoa	0,04197	0,417	84,703	6,433
44	trolebús	0,03975	0,341	85,044	5,263
45	Cercanías	0,03649	0,341	85,384	5,263

46	camioneta	0,03255	0,341	85,725	5,263
47	trineo	0,03235	0,379	86,104	5,848
48	parapente	0,02982	0,303	86,407	4,678
49	cohete espacial	0,02877	0,265	86,672	4,094
50	cuatro por cuatro	0,02831	0,227	86,899	3,509
51	todoterreno	0,02827	0,265	87,164	4,094
52	teleférico	0,02804	0,303	87,467	4,678
53	pie	0,02776	0,303	87,770	4,678
54	bote	0,02713	0,303	88,073	4,678
55	carro	0,02651	0,303	88,376	4,678
56	kayak	0,02557	0,265	88,641	4,094
57	nave espacial	0,02544	0,265	88,906	4,094
58	globo aerostático	0,02482	0,265	89,171	4,094
59	coche descapotable	0,02425	0,227	89,398	3,509
60	scooter	0,02311	0,189	89,587	2,924
61	mulo	0,02274	0,227	89,814	3,509
62	burro-taxi	0,02117	0,189	90,004	2,924
63	hidropedal	0,02007	0,227	90,231	3,509
64	funicular	0,01994	0,189	90,420	2,924
65	moto acuática	0,01874	0,227	90,647	3,509
66	patineta	0,01862	0,189	90,837	2,924
67	esquí	0,01751	0,227	91,064	3,509
68	paracaídas	0,01734	0,189	91,253	2,924
69	tren de alta velocidad	0,01714	0,151	91,405	2,339
70	zepelín	0,01670	0,189	91,594	2,924
71	carruaje	0,01667	0,189	91,783	2,924
72	Mini	0,01667	0,151	91,935	2,339
73	dirigible	0,01602	0,189	92,124	2,924
74	elefante	0,01500	0,189	92,314	2,924
75	góndola	0,01443	0,189	92,503	2,924
76	limusina	0,01361	0,151	92,654	2,339
77	ferrocarril	0,01315	0,151	92,806	2,339
78	tándem	0,01303	0,151	92,957	2,339
79	transbordador	0,01300	0,114	93,071	1,754
80	Seiscientos	0,01267	0,114	93,184	1,754
81	dromedario	0,01249	0,151	93,336	2,339
82	hidroavión	0,01206	0,114	93,449	1,754
83	carroza	0,01202	0,151	93,601	2,339
84	furgón	0,01162	0,114	93,715	1,754
85	asno	0,01122	0,114	93,828	1,754
86	autocaravana	0,01081	0,114	93,942	1,754
87	jet privado	0,01039	0,114	94,055	1,754
88	vehículo	0,01038	0,076	94,131	1,170
89	barco de vela	0,01008	0,114	94,245	1,754
90	carreta	0,01005	0,114	94,358	1,754
91	microbús	0,00968	0,114	94,472	1,754
92	carromato	0,00966	0,114	94,585	1,754
93	coche particular	0,00966	0,076	94,661	1,170

94	tabla de surf	0,00962	0,114	94,775	1,754
95	buey	0,00946	0,114	94,888	1,754
96	minibús	0,00943	0,114	95,002	1,754
97	calesa	0,00926	0,114	95,115	1,754
98	metro ligero	0,00881	0,076	95,191	1,170
99	monovolumen	0,00857	0,076	95,267	1,170
100	ala delta	0,00834	0,114	95,381	1,754
101	diligencia	0,00823	0,114	95,494	1,754
102	piragua	0,00816	0,114	95,608	1,754
103	tabla de snowboard	0,00813	0,114	95,721	1,754
104	cayuco	0,00808	0,114	95,835	1,754
105	buque	0,00798	0,076	95,911	1,170
106	catamarán	0,00797	0,076	95,986	1,170
107	automóvil	0,00782	0,076	96,062	1,170
108	potro	0,00768	0,076	96,138	1,170
109	coche deportivo	0,00763	0,076	96,214	1,170
110	carricoche	0,00744	0,076	96,289	1,170
111	ranchera	0,00732	0,076	96,365	1,170
112	ambulancia	0,00719	0,076	96,441	1,170
113	reno	0,00712	0,076	96,516	1,170
114	monorraíl	0,00710	0,076	96,592	1,170
115	montacargas	0,00710	0,076	96,668	1,170
116	ómnibus	0,00676	0,076	96,744	1,170
117	yegua	0,00674	0,076	96,819	1,170
118	barco de vapor	0,00670	0,076	96,895	1,170
119	ascensor	0,00652	0,076	96,971	1,170
120	aeropuerto	0,00647	0,076	97,047	1,170
121	telesilla	0,00599	0,076	97,122	1,170
122	zapato	0,00585	0,038	97,160	0,585
123	mula	0,00573	0,076	97,236	1,170
124	escalera mecánica	0,00564	0,076	97,312	1,170
125	andar	0,00549	0,076	97,387	1,170
126	portaaviones	0,00548	0,076	97,463	1,170
127	skate	0,00546	0,076	97,539	1,170
128	remolque	0,00507	0,076	97,615	1,170
129	transporte público	0,00506	0,038	97,652	0,585
130	vaporetto	0,00506	0,038	97,690	0,585
131	telecabina	0,00494	0,076	97,766	1,170
132	AT-AT	0,00483	0,038	97,804	0,585
133	tuk-tuk	0,00483	0,038	97,842	0,585
134	AT-ST	0,00460	0,038	97,880	0,585
135	Cuatro Latas	0,00460	0,038	97,917	0,585
136	locomotora	0,00460	0,038	97,955	0,585
137	berlina	0,00439	0,038	97,993	0,585
138	camión de bomberos	0,00439	0,038	98,031	0,585
139	planeador	0,00439	0,038	98,069	0,585
140	caminar	0,00418	0,038	98,107	0,585
141	ultraligero	0,00418	0,038	98,145	0,585

142	zódiac	0,00418	0,038	98,183	0,585
143	avión de caza	0,00362	0,038	98,220	0,585
144	balsa	0,00362	0,038	98,258	0,585
145	tren de mercancías	0,00362	0,038	98,296	0,585
146	vespa	0,00362	0,038	98,334	0,585
147	animal	0,00345	0,038	98,372	0,585
148	puerto	0,00345	0,038	98,410	0,585
149	carretilla	0,00329	0,038	98,448	0,585
150	estación de autobús	0,00329	0,038	98,485	0,585
151	noria	0,00329	0,038	98,523	0,585
152	tren eléctrico	0,00329	0,038	98,561	0,585
153	utilitario	0,00329	0,038	98,599	0,585
154	estación de tren	0,00314	0,038	98,637	0,585
155	jet	0,00314	0,038	98,675	0,585
156	lomo de burro	0,00314	0,038	98,713	0,585
157	poni	0,00314	0,038	98,750	0,585
158	embarcación	0,00299	0,038	98,788	0,585
159	esquí acuático	0,00299	0,038	98,826	0,585
160	estación de metro	0,00299	0,038	98,864	0,585
161	silla de ruedas	0,00299	0,038	98,902	0,585
162	coche de bebé	0,00285	0,038	98,940	0,585
163	coche teledirigido	0,00285	0,038	98,978	0,585
164	moto de agua	0,00285	0,038	99,016	0,585
165	pierna	0,00285	0,038	99,053	0,585
166	Renfe	0,00285	0,038	99,091	0,585
167	helipuerto	0,00272	0,038	99,129	0,585
168	máquina del tiempo	0,00272	0,038	99,167	0,585
169	transmediterráneo	0,00272	0,038	99,205	0,585
170	excavadora	0,00259	0,038	99,243	0,585
171	cohetes	0,00247	0,038	99,281	0,585
172	monociclo	0,00247	0,038	99,318	0,585
173	parada de autobús	0,00247	0,038	99,356	0,585
174	kilómetro	0,00235	0,038	99,394	0,585
175	carro de caballos	0,00224	0,038	99,432	0,585
176	viaje	0,00224	0,038	99,470	0,585
177	embarque	0,00214	0,038	99,508	0,585
178	(auto)bus turístico	0,00204	0,038	99,546	0,585
179	equipaje	0,00204	0,038	99,583	0,585
180	(auto)bus escolar	0,00194	0,038	99,621	0,585
181	facturar	0,00194	0,038	99,659	0,585
182	avestruz	0,00185	0,038	99,697	0,585
183	velocidad	0,00185	0,038	99,735	0,585
184	tráfico	0,00177	0,038	99,773	0,585
185	autovía	0,00168	0,038	99,811	0,585
186	autopista	0,00160	0,038	99,849	0,585
187	peaje	0,00153	0,038	99,886	0,585
188	peatón	0,00146	0,038	99,924	0,585
189	aparcamiento	0,00139	0,038	99,962	0,585
190	parking	0,00132	0,038	100,000	0,585

1.2.7. Juegos y distracciones

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	fútbol	0,47019	3,384	3,384	60,234
2	parchís	0,44286	3,055	6,439	54,386
3	baloncesto	0,36446	2,792	9,231	49,708
4	cine	0,35542	2,825	12,057	50,292
5	(juego de) cartas	0,31110	2,300	14,356	40,936
6	videojuego	0,27037	2,070	16,426	36,842
7	ajedrez	0,26909	1,971	18,397	35,088
8	tenis	0,24930	2,004	20,401	35,673
9	oca	0,24883	1,840	22,240	32,749
10	televisión	0,22784	1,807	24,047	32,164
11	leer	0,22765	1,840	25,887	32,749
12	deporte	0,21584	1,708	27,595	30,409
13	música	0,20842	1,675	29,271	29,825
14	teatro	0,18386	1,544	30,815	27,485
15	pasear	0,15688	1,380	32,194	24,561
16	bailar	0,15653	1,314	33,509	23,392
17	internet	0,15533	1,216	34,724	21,637
18	juego de mesa	0,14266	1,084	35,808	19,298
19	escondite	0,14053	1,150	36,958	20,468
20	pádel	0,13333	1,150	38,108	20,468
21	póquer	0,12705	0,953	39,060	16,959
22	damas	0,12673	1,051	40,112	18,713
23	Monopoly	0,12273	0,986	41,097	17,544
24	natación	0,12190	1,051	42,148	18,713
25	Trivial	0,11859	0,854	43,003	15,205
26	correr	0,11668	1,051	44,054	18,713
27	balonmano	0,11528	0,986	45,039	17,544
28	salir con amigos	0,10830	0,887	45,926	15,789
29	viajar	0,10444	0,887	46,813	15,789
30	pillapilla	0,09208	0,788	47,602	14,035
31	escuchar música	0,08758	0,723	48,325	12,865
32	senderismo	0,08258	0,756	49,080	13,450
33	ordenador	0,07938	0,657	49,737	11,696
34	comba	0,07718	0,624	50,361	11,111
35	patinar	0,07681	0,690	51,051	12,281
36	bicicleta	0,07540	0,690	51,741	12,281
37	cantar	0,07539	0,591	52,332	10,526
38	lectura	0,07517	0,591	52,924	10,526
39	petanca	0,07439	0,591	53,515	10,526
40	atletismo	0,07331	0,657	54,172	11,696
41	Play Station	0,07299	0,624	54,796	11,111
42	mus	0,07108	0,558	55,355	9,942
43	concierto	0,07018	0,624	55,979	11,111
44	película	0,06992	0,591	56,570	10,526

45	voleibol	0,06613	0,591	57,162	10,526
46	escribir	0,06546	0,591	57,753	10,526
47	pintar	0,06285	0,526	58,279	9,357
48	salir	0,06151	0,526	58,804	9,357
49	bolo	0,06019	0,493	59,297	8,772
50	dominó	0,05902	0,493	59,790	8,772
51	esquiar	0,05616	0,493	60,283	8,772
52	elástico	0,05532	0,427	60,710	7,602
53	libro	0,05449	0,427	61,137	7,602
54	Tabú	0,05422	0,526	61,662	9,357
55	ciclismo	0,05270	0,460	62,122	8,187
56	karaoke	0,05262	0,460	62,582	8,187
57	museo	0,05192	0,460	63,042	8,187
58	brisca	0,05074	0,361	63,403	6,433
59	rugby	0,04982	0,427	63,830	7,602
60	cinquillo	0,04969	0,394	64,225	7,018
61	dado	0,04954	0,394	64,619	7,018
62	discoteca	0,04943	0,427	65,046	7,602
63	golf	0,04797	0,394	65,440	7,018
64	serie	0,04599	0,361	65,802	6,433
65	Facebook	0,04584	0,394	66,196	7,018
66	playa	0,04356	0,394	66,590	7,018
67	pelota	0,04340	0,296	66,886	5,263
68	cocinar	0,04308	0,427	67,313	7,602
69	bar	0,04028	0,329	67,641	5,848
70	radio	0,03991	0,361	68,003	6,433
71	nadar	0,03931	0,361	68,364	6,433
72	tocar un instrumento	0,03919	0,296	68,660	5,263
73	chatear	0,03824	0,394	69,054	7,018
74	juego de ordenador	0,03818	0,263	69,317	4,678
75	videoconsola	0,03767	0,263	69,580	4,678
76	Scrabble	0,03619	0,296	69,875	5,263
77	crucigrama	0,03574	0,296	70,171	5,263
78	baile	0,03532	0,296	70,466	5,263
79	Cluedo	0,03486	0,263	70,729	4,678
80	gimnasio	0,03442	0,329	71,058	5,848
81	Wii	0,03433	0,296	71,353	5,263
82	billar	0,03404	0,263	71,616	4,678
83	gallinita ciega	0,03374	0,296	71,912	5,263
84	ir de compras	0,03267	0,296	72,208	5,263
85	tres en raya	0,03228	0,296	72,503	5,263
86	muñeco	0,03147	0,296	72,799	5,263
87	caminar	0,03140	0,329	73,127	5,848
88	escalada	0,03082	0,296	73,423	5,263
89	foto(grafía)	0,03060	0,263	73,686	4,678
90	bingo	0,02971	0,263	73,949	4,678
91	pintura	0,02920	0,230	74,179	4,094
92	Tuenti	0,02719	0,197	74,376	3,509

93	sudoku	0,02688	0,230	74,606	4,094
94	consola	0,02670	0,230	74,836	4,094
95	hockey	0,02656	0,230	75,066	4,094
96	mentiroso	0,02645	0,197	75,263	3,509
97	ping-pong	0,02637	0,263	75,526	4,678
98	dibujar	0,02620	0,197	75,723	3,509
99	restaurante	0,02538	0,197	75,920	3,509
100	red social	0,02503	0,164	76,084	2,924
101	Party & Co.	0,02492	0,230	76,314	4,094
102	juego de azar	0,02444	0,230	76,544	4,094
103	bádminton	0,02415	0,230	76,774	4,094
104	motociclismo	0,02397	0,197	76,971	3,509
105	rayuela	0,02368	0,197	77,168	3,509
106	puzle	0,02360	0,197	77,365	3,509
107	Tetris	0,02353	0,164	77,530	2,924
108	caballo	0,02315	0,197	77,727	3,509
109	charlar	0,02189	0,164	77,891	2,924
110	revista	0,02169	0,197	78,088	3,509
111	béisbol	0,02161	0,197	78,285	3,509
112	equitación	0,02103	0,197	78,482	3,509
113	spinning	0,02074	0,164	78,647	2,924
114	Messenger	0,02054	0,164	78,811	2,924
115	Scattergories	0,02048	0,164	78,975	2,924
116	solitario	0,02022	0,164	79,139	2,924
117	musical	0,01946	0,164	79,304	2,924
118	aerobic	0,01885	0,131	79,435	2,339
119	tocar la guitarra	0,01877	0,131	79,566	2,339
120	documental	0,01834	0,164	79,731	2,924
121	pub	0,01800	0,164	79,895	2,924
122	lotería	0,01799	0,164	80,059	2,924
123	step	0,01794	0,131	80,191	2,339
124	dardo	0,01791	0,164	80,355	2,924
125	tute	0,01785	0,131	80,486	2,339
126	beber	0,01763	0,164	80,650	2,924
127	escritura	0,01729	0,131	80,782	2,339
128	hablar	0,01683	0,164	80,946	2,924
129	yoga	0,01646	0,164	81,110	2,924
130	salir a cenar	0,01641	0,164	81,275	2,924
131	balón	0,01633	0,131	81,406	2,339
132	chat	0,01587	0,164	81,570	2,924
133	veo-veo	0,01582	0,164	81,735	2,924
134	esquí	0,01573	0,131	81,866	2,339
135	espectáculo	0,01573	0,131	81,997	2,339
136	Super Mario	0,01571	0,164	82,162	2,924
137	novela	0,01561	0,164	82,326	2,924
138	canicas	0,01544	0,131	82,457	2,339
139	salir de fiesta	0,01535	0,164	82,622	2,924
140	saltar	0,01511	0,131	82,753	2,339
141	coser	0,01492	0,131	82,884	2,339

142	chaquete	0,01469	0,099	82,983	1,754
143	footing	0,01454	0,131	83,114	2,339
144	Twitter	0,01435	0,131	83,246	2,339
145	snow(board)	0,01421	0,131	83,377	2,339
146	kárate	0,01418	0,131	83,509	2,339
147	navegar	0,01389	0,164	83,673	2,924
148	Pictionary	0,01375	0,099	83,771	1,754
149	escultura	0,01373	0,131	83,903	2,339
150	campo	0,01351	0,131	84,034	2,339
151	gimnasia	0,01345	0,131	84,166	2,339
152	SingStar	0,01329	0,099	84,264	1,754
153	alpinismo	0,01322	0,131	84,396	2,339
154	comer	0,01296	0,131	84,527	2,339
155	columpio	0,01277	0,099	84,625	1,754
156	excursión	0,01265	0,131	84,757	2,339
157	fútbolín	0,01261	0,099	84,855	1,754
158	yudo	0,01242	0,099	84,954	1,754
159	Youtube	0,01230	0,099	85,053	1,754
160	tocar el piano	0,01224	0,099	85,151	1,754
161	Risk	0,01220	0,099	85,250	1,754
162	deporte de riesgo	0,01197	0,131	85,381	2,339
163	ópera	0,01185	0,099	85,480	1,754
164	patinaje	0,01181	0,099	85,578	1,754
165	tomar café	0,01169	0,099	85,677	1,754
166	jugar	0,01162	0,099	85,775	1,754
167	parque de atracciones	0,01151	0,099	85,874	1,754
168	puenting	0,01145	0,131	86,005	2,339
169	viaje	0,01144	0,099	86,104	1,754
170	salir de marcha	0,01136	0,099	86,202	1,754
171	piso	0,01125	0,099	86,301	1,754
172	UNO	0,01095	0,131	86,432	2,339
173	Series Yonkis	0,01084	0,099	86,531	1,754
174	caza	0,01056	0,099	86,629	1,754
175	triatlón	0,01056	0,099	86,728	1,754
176	vela	0,01051	0,099	86,827	1,754
177	circo	0,01044	0,099	86,925	1,754
178	pesca	0,01027	0,099	87,024	1,754
179	Farm Ville	0,01014	0,066	87,089	1,170
180	Pilates	0,01003	0,099	87,188	1,754
181	bucear	0,00999	0,099	87,286	1,754
182	juego de naipes	0,00992	0,066	87,352	1,170
183	pollito inglés	0,00985	0,099	87,451	1,754
184	Barbie	0,00957	0,099	87,549	1,754
185	chinchón	0,00956	0,066	87,615	1,170
186	amigo	0,00951	0,099	87,714	1,754
187	cómic	0,00949	0,099	87,812	1,754
188	vídeo	0,00946	0,099	87,911	1,754
189	montañismo	0,00931	0,099	88,009	1,754

190	Game Boy	0,00922	0,131	88,141	2,339
191	arte	0,00911	0,066	88,206	1,170
192	policías y ladrones	0,00906	0,099	88,305	1,754
193	squash	0,00905	0,066	88,371	1,170
194	cricket	0,00900	0,099	88,469	1,754
195	estudiar	0,00880	0,099	88,568	1,754
196	cenar	0,00877	0,066	88,633	1,170
197	palmas	0,00863	0,099	88,732	1,754
198	cuento	0,00859	0,066	88,798	1,170
199	tomar el sol	0,00859	0,066	88,863	1,170
200	siesta	0,00857	0,066	88,929	1,170
201	carrera de coches	0,00842	0,099	89,028	1,754
202	Pro Evolution Soccer	0,00840	0,066	89,093	1,170
203	coche	0,00837	0,066	89,159	1,170
204	sopa de letras	0,00832	0,066	89,225	1,170
205	guitarra	0,00819	0,066	89,290	1,170
206	pescar	0,00819	0,066	89,356	1,170
207	obra	0,00801	0,066	89,422	1,170
208	barco	0,00789	0,066	89,488	1,170
209	carrera	0,00782	0,066	89,553	1,170
210	Playmobil	0,00781	0,099	89,652	1,754
211	salto de pértiga	0,00771	0,066	89,717	1,170
212	ir de tiendas	0,00768	0,066	89,783	1,170
213	Zelda	0,00768	0,066	89,849	1,170
214	patín	0,00768	0,066	89,915	1,170
215	escalar	0,00760	0,066	89,980	1,170
216	monumento	0,00754	0,066	90,046	1,170
217	ahorcado	0,00745	0,066	90,112	1,170
218	flauta	0,00744	0,066	90,177	1,170
219	paseo	0,00742	0,066	90,243	1,170
220	Nintendo	0,00738	0,066	90,309	1,170
221	waterpolo	0,00736	0,099	90,407	1,754
222	relevé	0,00735	0,066	90,473	1,170
223	gimnasia rítmica	0,00732	0,066	90,539	1,170
224	comprar	0,00730	0,099	90,637	1,754
225	danza	0,00728	0,066	90,703	1,170
226	tiro con arco	0,00725	0,066	90,769	1,170
227	andar	0,00718	0,066	90,834	1,170
228	Jungle Speed	0,00712	0,066	90,900	1,170
229	Hotmail	0,00709	0,066	90,966	1,170
230	salsa	0,00703	0,066	91,032	1,170
231	DVD	0,00694	0,066	91,097	1,170
232	piedra papel o tijera	0,00691	0,066	91,163	1,170
233	turismo	0,00687	0,099	91,261	1,754
234	Xbox	0,00686	0,066	91,327	1,170
235	balonvolea	0,00677	0,066	91,393	1,170
236	piragüismo	0,00677	0,066	91,459	1,170
237	pares o nones	0,00659	0,066	91,524	1,170

238	submarinismo	0,00658	0,066	91,590	1,170
239	telenovela	0,00658	0,066	91,656	1,170
240	móvil	0,00644	0,099	91,754	1,754
241	ruleta	0,00639	0,066	91,820	1,170
242	dibujo	0,00634	0,066	91,886	1,170
243	pasear al perro	0,00630	0,066	91,951	1,170
244	PSP	0,00628	0,066	92,017	1,170
245	ballet	0,00615	0,066	92,083	1,170
246	test	0,00614	0,066	92,148	1,170
247	paleta	0,00612	0,066	92,214	1,170
248	decorar	0,00609	0,066	92,280	1,170
249	surfear	0,00599	0,066	92,346	1,170
250	(máquina) tragaperras	0,00592	0,066	92,411	1,170
251	máster	0,00585	0,033	92,444	0,585
252	palomita	0,00585	0,033	92,477	0,585
253	Rummikub	0,00585	0,033	92,510	0,585
254	mascota	0,00576	0,066	92,576	1,170
255	punto de cruz	0,00571	0,066	92,641	1,170
256	Buzz!	0,00557	0,033	92,674	0,585
257	el patio de mi casa	0,00557	0,033	92,707	0,585
258	partido	0,00557	0,033	92,740	0,585
259	rol	0,00557	0,033	92,773	0,585
260	tiovivo	0,00557	0,033	92,806	0,585
261	toros	0,00557	0,033	92,838	0,585
262	tres en línea	0,00557	0,033	92,871	0,585
263	yoyó	0,00557	0,033	92,904	0,585
264	conducir	0,00557	0,066	92,970	1,170
265	hablar por teléfono	0,00543	0,066	93,035	1,170
266	alcohol	0,00531	0,033	93,068	0,585
267	naturaleza	0,00531	0,033	93,101	0,585
268	trekking	0,00531	0,033	93,134	0,585
269	corro de la patata	0,00531	0,066	93,200	1,170
270	pensar	0,00531	0,066	93,265	1,170
271	monopatín	0,00527	0,066	93,331	1,170
272	esgrima	0,00514	0,066	93,397	1,170
273	acampar	0,00506	0,033	93,430	0,585
274	carambola	0,00506	0,033	93,463	0,585
275	carricoche	0,00506	0,033	93,495	0,585
276	fitness	0,00506	0,033	93,528	0,585
277	juego de rol	0,00506	0,033	93,561	0,585
278	juego on line	0,00506	0,033	93,594	0,585
279	noticia	0,00506	0,033	93,627	0,585
280	Pet Society	0,00506	0,033	93,660	0,585
281	canasta	0,00483	0,033	93,693	0,585
282	colorear	0,00483	0,033	93,725	0,585
283	freaky	0,00483	0,033	93,758	0,585

284	hijo puta	0,00483	0,033	93,791	0,585
285	jeroglífico	0,00483	0,033	93,824	0,585
286	juego	0,00483	0,033	93,857	0,585
287	rocódromo	0,00483	0,033	93,890	0,585
288	ruleta rusa	0,00483	0,033	93,922	0,585
289	sexo	0,00483	0,033	93,955	0,585
290	kickboxing	0,00473	0,066	94,021	1,170
291	conejo de la suerte	0,00471	0,066	94,087	1,170
292	coleccionar	0,00460	0,033	94,120	0,585
293	decoración	0,00460	0,033	94,152	0,585
294	dibujo animado	0,00460	0,033	94,185	0,585
295	escalerilla	0,00460	0,033	94,218	0,585
296	patinete	0,00460	0,033	94,251	0,585
297	poli-ladro	0,00460	0,033	94,284	0,585
298	siete diferencias	0,00460	0,033	94,317	0,585
299	Sims	0,00460	0,033	94,350	0,585
300	trineo	0,00460	0,033	94,382	0,585
301	emborracharse	0,00445	0,066	94,448	1,170
302	¿Dónde está Wally?	0,00439	0,033	94,481	0,585
303	botellón	0,00439	0,033	94,514	0,585
304	coleccionismo	0,00439	0,033	94,547	0,585
305	esculpir	0,00439	0,033	94,580	0,585
306	esquí de fondo	0,00439	0,033	94,612	0,585
307	fútbol sala	0,00439	0,033	94,645	0,585
308	galería de arte	0,00439	0,033	94,678	0,585
309	noria	0,00439	0,033	94,711	0,585
310	rey del tres	0,00439	0,033	94,744	0,585
311	periódico	0,00432	0,066	94,809	1,170
312	baraja	0,00418	0,033	94,842	0,585
313	bordar	0,00418	0,033	94,875	0,585
314	bricolaje	0,00418	0,033	94,908	0,585
315	coche teledirigido	0,00418	0,033	94,941	0,585
316	esquí alpino	0,00418	0,033	94,974	0,585
317	explorar	0,00418	0,033	95,007	0,585
318	informativo	0,00418	0,033	95,039	0,585
319	juego de magia	0,00418	0,033	95,072	0,585
320	palacio	0,00418	0,033	95,105	0,585
321	serie de televisión	0,00418	0,033	95,138	0,585
322	violín	0,00418	0,033	95,171	0,585
323	bicicleta elíptica	0,00399	0,033	95,204	0,585
324	boxeo	0,00399	0,033	95,237	0,585
325	buceo	0,00399	0,033	95,269	0,585
326	burro	0,00399	0,033	95,302	0,585
327	cafetería	0,00399	0,033	95,335	0,585
328	catedral	0,00399	0,033	95,368	0,585
329	cometa	0,00399	0,033	95,401	0,585

330	componer música	0,00399	0,033	95,434	0,585
331	eslalon	0,00399	0,033	95,466	0,585
332	ficción	0,00399	0,033	95,499	0,585
333	hacer cine	0,00399	0,033	95,532	0,585
334	hermano	0,00399	0,033	95,565	0,585
335	investigar	0,00399	0,033	95,598	0,585
336	juego didáctico	0,00399	0,033	95,631	0,585
337	ronda	0,00399	0,033	95,664	0,585
338	tobogán	0,00399	0,033	95,696	0,585
339	astronomía	0,00380	0,033	95,729	0,585
340	bicicleta estática	0,00380	0,033	95,762	0,585
341	El Gordo	0,00380	0,033	95,795	0,585
342	halterofilia	0,00380	0,033	95,828	0,585
343	iglesia	0,00380	0,033	95,861	0,585
344	limpiar la casa	0,00380	0,033	95,894	0,585
345	mando a distancia	0,00380	0,033	95,926	0,585
346	parque	0,00380	0,033	95,959	0,585
347	quiniela	0,00380	0,033	95,992	0,585
348	tambor	0,00380	0,033	96,025	0,585
349	tertulia	0,00380	0,033	96,058	0,585
350	tirar piedras al tren	0,00380	0,033	96,091	0,585
351	¿Quién es quién?	0,00362	0,033	96,124	0,585
352	apuesta	0,00362	0,033	96,156	0,585
353	automovilismo	0,00362	0,033	96,189	0,585
354	bolera	0,00362	0,033	96,222	0,585
355	crear	0,00362	0,033	96,255	0,585
356	masturbar	0,00362	0,033	96,288	0,585
357	minigolf	0,00362	0,033	96,321	0,585
358	peluquería	0,00362	0,033	96,353	0,585
359	piano	0,00362	0,033	96,386	0,585
360	surf	0,00362	0,033	96,419	0,585
361	Wikipedia	0,00362	0,033	96,452	0,585
362	World of Warcraft	0,00362	0,033	96,485	0,585
363	competición	0,00345	0,033	96,518	0,585
364	descenso	0,00345	0,033	96,551	0,585
365	dormir	0,00345	0,033	96,583	0,585
366	estiramiento	0,00345	0,033	96,616	0,585
367	Google	0,00345	0,033	96,649	0,585
368	médicos	0,00345	0,033	96,682	0,585
369	musculación	0,00345	0,033	96,715	0,585
370	rafting	0,00345	0,033	96,748	0,585
371	taichí	0,00345	0,033	96,781	0,585
372	gastronomía	0,00329	0,033	96,813	0,585
373	Ikea	0,00329	0,033	96,846	0,585
374	ola	0,00329	0,033	96,879	0,585
375	pulso	0,00329	0,033	96,912	0,585

376	salto de la burra	0,00329	0,033	96,945	0,585
377	churro pico o terna	0,00314	0,033	96,978	0,585
378	Dicciopinta	0,00314	0,033	97,011	0,585
379	espeleología	0,00314	0,033	97,043	0,585
380	feria	0,00314	0,033	97,076	0,585
381	juego del chupito	0,00314	0,033	97,109	0,585
382	pulso chino	0,00314	0,033	97,142	0,585
383	Sonic	0,00314	0,033	97,175	0,585
384	tapersex	0,00314	0,033	97,208	0,585
385	tejer	0,00314	0,033	97,240	0,585
386	blog	0,00299	0,033	97,273	0,585
387	cazar	0,00299	0,033	97,306	0,585
388	CD	0,00299	0,033	97,339	0,585
389	clase	0,00299	0,033	97,372	0,585
390	compra	0,00299	0,033	97,405	0,585
391	cuatro en raya	0,00299	0,033	97,438	0,585
392	jogging	0,00299	0,033	97,470	0,585
393	monólogo	0,00299	0,033	97,503	0,585
394	ocio nocturno	0,00299	0,033	97,536	0,585
395	reunirse con amigos	0,00299	0,033	97,569	0,585
396	salto de altura	0,00299	0,033	97,602	0,585
397	salto de longitud	0,00299	0,033	97,635	0,585
398	aquagym	0,00285	0,033	97,668	0,585
399	cabalgar	0,00285	0,033	97,700	0,585
400	camping	0,00285	0,033	97,733	0,585
401	casitas	0,00285	0,033	97,766	0,585
402	deporte de contacto	0,00285	0,033	97,799	0,585
403	foro	0,00285	0,033	97,832	0,585
404	fútbol americano	0,00285	0,033	97,865	0,585
405	pasear por el campo	0,00285	0,033	97,898	0,585
406	programa	0,00285	0,033	97,930	0,585
407	recortable	0,00285	0,033	97,963	0,585
408	soldado	0,00285	0,033	97,996	0,585
409	tienda	0,00285	0,033	98,029	0,585
410	animal	0,00272	0,033	98,062	0,585
411	apostar	0,00272	0,033	98,095	0,585
412	balón prisionero	0,00272	0,033	98,127	0,585
413	carrera de caballos	0,00272	0,033	98,160	0,585
414	disfraz	0,00272	0,033	98,193	0,585
415	funky	0,00272	0,033	98,226	0,585
416	intercambio de idiomas	0,00272	0,033	98,259	0,585
417	Pokemon	0,00272	0,033	98,292	0,585
418	taba	0,00272	0,033	98,325	0,585
419	trompo	0,00272	0,033	98,357	0,585
420	aprender	0,00259	0,033	98,390	0,585

421	aro	0,00259	0,033	98,423	0,585
422	ayudar	0,00259	0,033	98,456	0,585
423	balconing	0,00259	0,033	98,489	0,585
424	Buscaminas	0,00259	0,033	98,522	0,585
425	caballería	0,00259	0,033	98,555	0,585
426	divertirse	0,00259	0,033	98,587	0,585
427	flamenco	0,00259	0,033	98,620	0,585
428	fútbol playa	0,00259	0,033	98,653	0,585
429	peonza	0,00259	0,033	98,686	0,585
430	Super Pang	0,00259	0,033	98,719	0,585
431	Bubble Bobble	0,00247	0,033	98,752	0,585
432	canoa	0,00247	0,033	98,784	0,585
433	comecocos	0,00247	0,033	98,817	0,585
434	deporte de aventura	0,00247	0,033	98,850	0,585
435	diario	0,00247	0,033	98,883	0,585
436	Nancy	0,00247	0,033	98,916	0,585
437	reguetón	0,00247	0,033	98,949	0,585
438	sauna	0,00247	0,033	98,982	0,585
439	voluntariado	0,00247	0,033	99,014	0,585
440	chapas	0,00235	0,033	99,047	0,585
441	culebrón	0,00235	0,033	99,080	0,585
442	excursión a la naturaleza	0,00235	0,033	99,113	0,585
443	idioma	0,00235	0,033	99,146	0,585
444	kayak	0,00235	0,033	99,179	0,585
445	meditación	0,00235	0,033	99,212	0,585
446	paintball	0,00235	0,033	99,244	0,585
447	pesas	0,00235	0,033	99,277	0,585
448	tenis de mesa	0,00235	0,033	99,310	0,585
449	conversar	0,00224	0,033	99,343	0,585
450	exposición	0,00224	0,033	99,376	0,585
451	gesto	0,00224	0,033	99,409	0,585
452	maratón	0,00224	0,033	99,442	0,585
453	paracaídas	0,00224	0,033	99,474	0,585
454	patinaje sobre hielo	0,00224	0,033	99,507	0,585
455	pelota vasca	0,00224	0,033	99,540	0,585
456	Sega Mega Drive	0,00224	0,033	99,573	0,585
457	zapatilla por detrás	0,00224	0,033	99,606	0,585
458	taekwondo	0,00214	0,033	99,639	0,585
459	Yathzee	0,00214	0,033	99,671	0,585
460	hablar por WhatsApp	0,00204	0,033	99,704	0,585
461	subir al monte	0,00204	0,033	99,737	0,585
462	carrera de sacos	0,00194	0,033	99,770	0,585
463	esquí acuático	0,00194	0,033	99,803	0,585
464	pasear animales de compañía	0,00194	0,033	99,836	0,585

465	psicólogo	0,00185	0,033	99,869	0,585
466	canasta revuelta	0,00177	0,033	99,901	0,585
467	hípica	0,00177	0,033	99,934	0,585
468	fumar porros	0,00153	0,033	99,967	0,585
469	patio de mi casa	0,00153	0,033	100,000	0,585

1.2.8. Profesiones y oficios

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	profesor	0,78858	4,306	4,306	90,643
2	médico	0,46564	2,972	7,278	62,573
3	enfermero	0,31098	2,222	9,500	46,784
4	policía	0,29559	2,167	11,667	45,614
5	abogado	0,29363	2,000	13,667	42,105
6	ingeniero	0,27354	1,778	15,444	37,427
7	maestro	0,26079	1,556	17,000	32,749
8	camarero	0,24494	1,833	18,833	38,596
9	arquitecto	0,22886	1,528	20,361	32,164
10	cocinero	0,20936	1,639	22,000	34,503
11	fontanero	0,19821	1,361	23,361	28,655
12	bombero	0,19170	1,472	24,833	30,994
13	panadero	0,18823	1,250	26,083	26,316
14	conductor	0,18248	1,417	27,500	29,825
15	albañil	0,18057	1,278	28,778	26,901
16	pintor	0,17975	1,222	30,000	25,731
17	electricista	0,17552	1,333	31,333	28,070
18	taxista	0,15530	1,167	32,500	24,561
19	actor	0,14190	1,083	33,583	22,807
20	carpintero	0,13849	0,944	34,528	19,883
21	traductor	0,13550	0,917	35,444	19,298
22	doctor	0,13465	0,917	36,361	19,298
23	limpiador	0,13249	1,083	37,444	22,807
24	dependiente	0,12836	1,111	38,556	23,392
25	carnicero	0,11877	0,861	39,417	18,129
26	músico	0,11478	0,750	40,167	15,789
27	informático	0,11337	0,833	41,000	17,544
28	administrativo	0,11288	0,861	41,861	18,129
29	juez	0,11092	0,833	42,694	17,544
30	barrendero	0,11079	0,833	43,528	17,544
31	político	0,11035	0,889	44,417	18,713
32	psicólogo	0,10662	0,806	45,222	16,959
33	periodista	0,10589	0,833	46,056	17,544
34	director	0,10301	0,889	46,944	18,713
35	jardinero	0,09557	0,778	47,722	16,374
36	conserje	0,09552	0,750	48,472	15,789
37	peluquero	0,09481	0,639	49,111	13,450
38	cantante	0,09050	0,778	49,889	16,374
39	banquero	0,08275	0,694	50,583	14,620
40	intérprete	0,08266	0,583	51,167	12,281
41	agricultor	0,08211	0,667	51,833	14,035
42	secretario	0,08128	0,667	52,500	14,035
43	pescadero	0,08091	0,667	53,167	14,035
44	empresario	0,07924	0,694	53,861	14,620
45	escritor	0,07837	0,583	54,444	12,281

46	futbolista	0,07764	0,611	55,056	12,865
47	artista	0,07536	0,528	55,583	11,111
48	piloto	0,07158	0,611	56,194	12,865
49	vendedor	0,07132	0,556	56,750	11,696
50	estudiante	0,06981	0,556	57,306	11,696
51	ama de casa	0,06896	0,528	57,833	11,111
52	diseñador	0,06792	0,556	58,389	11,696
53	dentista	0,06668	0,528	58,917	11,111
54	frutero	0,06628	0,556	59,472	11,696
55	mecánico	0,06596	0,444	59,917	9,357
56	biólogo	0,06534	0,583	60,500	12,281
57	camionero	0,06509	0,556	61,056	11,696
58	químico	0,06470	0,472	61,528	9,942
59	investigador	0,06455	0,417	61,944	8,772
60	cajero	0,06039	0,556	62,500	11,696
61	farmacéutico	0,05934	0,500	63,000	10,526
62	azafato	0,05874	0,444	63,444	9,357
63	recepcionista	0,05737	0,472	63,917	9,942
64	presentador	0,05614	0,500	64,417	10,526
65	celador	0,05465	0,444	64,861	9,357
66	constructor	0,05309	0,361	65,222	7,602
67	cartero	0,05207	0,306	65,528	6,433
68	guía turístico	0,05137	0,361	65,889	7,602
69	basurero	0,04951	0,333	66,222	7,018
70	tendero	0,04939	0,361	66,583	7,602
71	funcionario	0,04831	0,361	66,944	7,602
72	filólogo	0,04817	0,361	67,306	7,602
73	Guardia Civil	0,04509	0,361	67,667	7,602
74	economista	0,04436	0,361	68,028	7,602
75	matemático	0,04363	0,361	68,389	7,602
76	chófer	0,04301	0,389	68,778	8,187
77	militar	0,04301	0,333	69,111	7,018
78	modelo	0,04131	0,333	69,444	7,018
79	peón	0,04128	0,333	69,778	7,018
80	alcalde	0,03910	0,306	70,083	6,433
81	comercial	0,03905	0,361	70,444	7,602
82	escultor	0,03902	0,278	70,722	5,848
83	psiquiatra	0,03858	0,306	71,028	6,433
84	bibliotecario	0,03852	0,278	71,306	5,848
85	contable	0,03701	0,278	71,583	5,848
86	ganadero	0,03680	0,278	71,861	5,848
87	historiador	0,03669	0,278	72,139	5,848
88	notario	0,03647	0,278	72,417	5,848
89	torero	0,03613	0,278	72,694	5,848
90	charcutero	0,03563	0,250	72,944	5,263
91	arqueólogo	0,03438	0,250	73,194	5,263
92	ginecólogo	0,03412	0,278	73,472	5,848
93	físico	0,03406	0,278	73,750	5,848
94	bailarín	0,03235	0,333	74,083	7,018

95	obrero	0,03228	0,222	74,306	4,678
96	aparejador	0,03186	0,250	74,556	5,263
97	docente	0,03157	0,167	74,722	3,509
98	guitarrista	0,03111	0,250	74,972	5,263
99	fotógrafo	0,03042	0,222	75,194	4,678
100	zapatero	0,02940	0,250	75,444	5,263
101	sastre	0,02937	0,222	75,667	4,678
102	deportista	0,02936	0,222	75,889	4,678
103	pedagogo	0,02932	0,167	76,056	3,509
104	pescador	0,02918	0,250	76,306	5,263
105	jefe	0,02896	0,222	76,528	4,678
106	decorador	0,02739	0,222	76,750	4,678
107	asistente	0,02728	0,194	76,944	4,094
108	pastelero	0,02609	0,194	77,139	4,094
109	marinero	0,02598	0,194	77,333	4,094
110	herrero	0,02575	0,194	77,528	4,094
111	oficinista	0,02572	0,194	77,722	4,094
112	encofrador	0,02534	0,194	77,917	4,094
113	gerente	0,02464	0,222	78,139	4,678
114	ATS	0,02459	0,222	78,361	4,678
115	parado	0,02306	0,139	78,500	2,924
116	teleoperador	0,02306	0,194	78,694	4,094
117	pediatra	0,02293	0,167	78,861	3,509
118	costurero	0,02259	0,194	79,056	4,094
119	publicista	0,02251	0,139	79,194	2,924
120	fiscal	0,02213	0,167	79,361	3,509
121	programador	0,02213	0,167	79,528	3,509
122	guardaespaldas	0,02207	0,194	79,722	4,094
123	cirujano	0,02194	0,167	79,889	3,509
124	violinista	0,02189	0,139	80,028	2,924
125	gobernanta	0,02159	0,194	80,222	4,094
126	pianista	0,02148	0,167	80,389	3,509
127	autónomo	0,02143	0,194	80,583	4,094
128	óptico	0,02135	0,139	80,722	2,924
129	guardia de seguridad	0,02131	0,167	80,889	3,509
130	científico	0,02118	0,139	81,028	2,924
131	cerrajero	0,02056	0,167	81,194	3,509
132	bróker	0,02053	0,139	81,333	2,924
133	soldado	0,01990	0,139	81,472	2,924
134	granjero	0,01938	0,167	81,639	3,509
135	asesor	0,01925	0,167	81,806	3,509
136	astronauta	0,01914	0,139	81,944	2,924
137	repartidor	0,01881	0,167	82,111	3,509
138	fisioterapeuta	0,01852	0,194	82,306	4,094
139	actriz	0,01850	0,111	82,417	2,339
140	prostituto	0,01822	0,139	82,556	2,924
141	ministro	0,01793	0,139	82,694	2,924
142	reponedor	0,01761	0,194	82,889	4,094
143	director de cine	0,01754	0,139	83,028	2,924

144	niñera	0,01744	0,167	83,194	3,509
145	soldador	0,01743	0,139	83,333	2,924
146	canguro	0,01728	0,139	83,472	2,924
147	empleado de banco	0,01724	0,139	83,611	2,924
148	presidente	0,01699	0,167	83,778	3,509
149	diputado	0,01697	0,111	83,889	2,339
150	procurador	0,01671	0,111	84,000	2,339
151	topógrafo	0,01667	0,111	84,111	2,339
152	ladrón	0,01648	0,111	84,222	2,339
153	perito	0,01647	0,139	84,361	2,924
154	veterinario	0,01594	0,111	84,472	2,339
155	maquinista	0,01579	0,139	84,611	2,924
156	filósofo	0,01573	0,139	84,750	2,924
157	director de orquesta	0,01571	0,111	84,861	2,339
158	cabrero	0,01565	0,167	85,028	3,509
159	monitor	0,01563	0,139	85,167	2,924
160	chef	0,01532	0,111	85,278	2,339
161	jugador de baloncesto	0,01516	0,139	85,417	2,924
162	comerciante	0,01513	0,111	85,528	2,339
163	cámara	0,01495	0,111	85,639	2,339
164	enseñante	0,01487	0,083	85,722	1,754
165	productor	0,01481	0,111	85,833	2,339
166	enterrador	0,01441	0,111	85,944	2,339
167	becario	0,01432	0,083	86,028	1,754
168	delineante	0,01432	0,111	86,139	2,339
169	poeta	0,01416	0,111	86,250	2,339
170	cantautor	0,01377	0,111	86,361	2,339
171	pinche	0,01357	0,111	86,472	2,339
172	lingüista	0,01336	0,083	86,556	1,754
173	concejal	0,01324	0,111	86,667	2,339
174	florista	0,01324	0,111	86,778	2,339
175	dibujante	0,01295	0,083	86,861	1,754
176	jefe de obra	0,01285	0,111	86,972	2,339
177	anestesista	0,01272	0,083	87,056	1,754
178	auxiliar de enfermería	0,01255	0,111	87,167	2,339
179	vigilante	0,01246	0,111	87,278	2,339
180	apicultor	0,01246	0,111	87,389	2,339
181	quiosquero	0,01244	0,111	87,500	2,339
182	orientador	0,01222	0,083	87,583	1,754
183	taxidermista	0,01214	0,111	87,694	2,339
184	conductor de autobús	0,01210	0,111	87,806	2,339
185	agente de viajes	0,01202	0,083	87,889	1,754
186	sumiller	0,01177	0,083	87,972	1,754
187	inspector	0,01177	0,083	88,056	1,754
188	arquitecto técnico	0,01170	0,056	88,111	1,170
189	esteticista	0,01144	0,083	88,194	1,754

190	controlador aéreo	0,01139	0,111	88,306	2,339
191	transportista	0,01125	0,111	88,417	2,339
192	guionista	0,01123	0,083	88,500	1,754
193	tutor	0,01115	0,056	88,556	1,170
194	geógrafo	0,01075	0,083	88,639	1,754
195	interiorista	0,01036	0,083	88,722	1,754
196	guardia	0,01013	0,083	88,806	1,754
197	auxiliar	0,01011	0,083	88,889	1,754
198	traumatólogo	0,01006	0,083	88,972	1,754
199	técnico	0,01002	0,083	89,056	1,754
200	portero	0,00994	0,111	89,167	2,339
201	chapista	0,00971	0,083	89,250	1,754
202	payaso	0,00967	0,111	89,361	2,339
203	gestor	0,00959	0,083	89,444	1,754
204	auxiliar de clínica	0,00955	0,083	89,528	1,754
205	socorrista	0,00952	0,083	89,611	1,754
206	oculista	0,00948	0,083	89,694	1,754
207	pastor	0,00937	0,083	89,778	1,754
208	atleta	0,00932	0,083	89,861	1,754
209	cura	0,00931	0,083	89,944	1,754
210	cardiólogo	0,00929	0,083	90,028	1,754
211	telefonista	0,00927	0,083	90,111	1,754
212	animador	0,00921	0,056	90,167	1,170
213	top-model	0,00914	0,083	90,250	1,754
214	librero	0,00914	0,083	90,333	1,754
215	confitero	0,00899	0,056	90,389	1,170
216	monja	0,00888	0,083	90,472	1,754
217	educador social	0,00878	0,056	90,528	1,170
218	capataz	0,00877	0,056	90,583	1,170
219	podólogo	0,00876	0,083	90,667	1,754
220	botones	0,00876	0,083	90,750	1,754
221	mago	0,00861	0,083	90,833	1,754
222	compositor	0,00857	0,083	90,917	1,754
223	cantaor	0,00851	0,083	91,000	1,754
224	mercenario	0,00845	0,056	91,056	1,170
225	analista	0,00845	0,056	91,111	1,170
226	trompetista	0,00840	0,056	91,167	1,170
227	encargado	0,00836	0,056	91,222	1,170
228	neurólogo	0,00822	0,056	91,278	1,170
229	educador	0,00810	0,083	91,361	1,754
230	chatarrero	0,00805	0,056	91,417	1,170
231	consultor	0,00805	0,056	91,472	1,170
232	entrenador	0,00805	0,056	91,528	1,170
233	masajista	0,00804	0,083	91,611	1,754
234	director de banco	0,00803	0,056	91,667	1,170
235	gruista	0,00798	0,056	91,722	1,170
236	consejero	0,00797	0,056	91,778	1,170
237	contratista	0,00796	0,056	91,833	1,170
238	tenista	0,00792	0,083	91,917	1,754

239	interventor	0,00789	0,056	91,972	1,170
240	sicario	0,00771	0,056	92,028	1,170
241	jeque árabe	0,00768	0,056	92,083	1,170
242	joyero	0,00765	0,056	92,139	1,170
243	odontólogo	0,00760	0,056	92,194	1,170
244	jefe de mantenimiento	0,00744	0,056	92,250	1,170
245	verdulero	0,00743	0,083	92,333	1,754
246	restaurador	0,00742	0,056	92,389	1,170
247	campesino	0,00741	0,083	92,472	1,754
248	acomodador	0,00732	0,083	92,556	1,754
249	minero	0,00717	0,056	92,611	1,170
250	magistrado	0,00701	0,056	92,667	1,170
251	aguador	0,00690	0,056	92,722	1,170
252	lechero	0,00677	0,056	92,778	1,170
253	carcelero	0,00670	0,056	92,833	1,170
254	especulador	0,00668	0,056	92,889	1,170
255	administrador	0,00664	0,056	92,944	1,170
256	bailaor	0,00659	0,056	93,000	1,170
257	estilista	0,00658	0,056	93,056	1,170
258	serviente	0,00658	0,056	93,111	1,170
259	afilador	0,00652	0,056	93,167	1,170
260	peón de albañil	0,00644	0,056	93,222	1,170
261	mánager	0,00628	0,056	93,278	1,170
262	timador	0,00627	0,056	93,333	1,170
263	aviador	0,00612	0,056	93,389	1,170
264	coronel	0,00604	0,056	93,444	1,170
265	locutor	0,00604	0,056	93,500	1,170
266	médico de cabecera	0,00604	0,056	93,556	1,170
267	ejecutivo	0,00594	0,056	93,611	1,170
268	ecologista	0,00585	0,028	93,639	0,585
269	trabajo	0,00585	0,028	93,667	0,585
270	agente inmobiliario	0,00584	0,056	93,722	1,170
271	agente secreto	0,00557	0,028	93,750	0,585
272	ingeniero químico	0,00557	0,028	93,778	0,585
273	paro	0,00557	0,028	93,806	0,585
274	violonchelista	0,00557	0,028	93,833	0,585
275	espía	0,00557	0,056	93,889	1,170
276	butanero	0,00546	0,056	93,944	1,170
277	coreógrafo	0,00534	0,056	94,000	1,170
278	revisor	0,00532	0,056	94,056	1,170
279	actor porno	0,00531	0,028	94,083	0,585
280	asistenta	0,00531	0,028	94,111	0,585
281	au pair	0,00531	0,028	94,139	0,585
282	estadista	0,00531	0,028	94,167	0,585
283	ingeniero eléctrico	0,00531	0,028	94,194	0,585

284	tornero	0,00531	0,028	94,222	0,585
285	jornalero	0,00523	0,056	94,278	1,170
286	auxiliar administrativo	0,00520	0,056	94,333	1,170
287	clarinetista	0,00506	0,028	94,361	0,585
288	deshollinador	0,00506	0,028	94,389	0,585
289	detective	0,00506	0,028	94,417	0,585
290	fresador	0,00506	0,028	94,444	0,585
291	ingeniero industrial	0,00506	0,028	94,472	0,585
292	ingeniero mecánico	0,00506	0,028	94,500	0,585
293	maître	0,00506	0,028	94,528	0,585
294	terminólogo	0,00506	0,028	94,556	0,585
295	ornitólogo	0,00499	0,056	94,611	1,170
296	árbitro	0,00496	0,056	94,667	1,170
297	ama de llaves	0,00483	0,056	94,722	1,170
298	camarero de pisos	0,00483	0,028	94,750	0,585
299	catador	0,00483	0,028	94,778	0,585
300	diplomático	0,00483	0,028	94,806	0,585
301	empleado	0,00483	0,028	94,833	0,585
302	enólogo	0,00483	0,028	94,861	0,585
303	ingeniero informático	0,00483	0,028	94,889	0,585
304	matrón	0,00483	0,028	94,917	0,585
305	operario	0,00483	0,028	94,944	0,585
306	traductor jurado	0,00483	0,028	94,972	0,585
307	catedrático	0,00471	0,056	95,028	1,170
308	cristalero	0,00461	0,056	95,083	1,170
309	forense	0,00460	0,028	95,111	0,585
310	ingeniero de sistemas	0,00460	0,028	95,139	0,585
311	profesor de secundaria	0,00460	0,028	95,167	0,585
312	realizador	0,00460	0,028	95,194	0,585
313	verdugo	0,00460	0,028	95,222	0,585
314	yesero	0,00460	0,028	95,250	0,585
315	auditor	0,00439	0,028	95,278	0,585
316	creativo	0,00439	0,028	95,306	0,585
317	diseñador gráfico	0,00439	0,028	95,333	0,585
318	endocrino	0,00439	0,028	95,361	0,585
319	enfoscador	0,00439	0,028	95,389	0,585
320	francotirador	0,00439	0,028	95,417	0,585
321	friegaplatos	0,00439	0,028	95,444	0,585
322	funámbulo	0,00439	0,028	95,472	0,585
323	herrador	0,00439	0,028	95,500	0,585
324	logopeda	0,00439	0,028	95,528	0,585
325	maestro de escuela	0,00439	0,028	95,556	0,585
326	peletero	0,00439	0,028	95,583	0,585

327	vinicultor	0,00439	0,028	95,611	0,585
328	visitador médico	0,00438	0,056	95,667	1,170
329	modisto	0,00428	0,056	95,722	1,170
330	caudillo	0,00418	0,028	95,750	0,585
331	chacha	0,00418	0,028	95,778	0,585
332	eurodiputado	0,00418	0,028	95,806	0,585
333	heladero	0,00418	0,028	95,833	0,585
334	hotelero	0,00418	0,028	95,861	0,585
335	ingeniero de caminos	0,00418	0,028	95,889	0,585
336	jubilado	0,00418	0,028	95,917	0,585
337	otorrino	0,00418	0,028	95,944	0,585
338	profesor universitario	0,00418	0,028	95,972	0,585
339	tejedor	0,00418	0,028	96,000	0,585
340	marino	0,00408	0,056	96,056	1,170
341	maquillador	0,00399	0,056	96,111	1,170
342	dermatólogo	0,00399	0,028	96,139	0,585
343	mancebo	0,00399	0,028	96,167	0,585
344	operador de cámara	0,00399	0,028	96,194	0,585
345	sereno	0,00399	0,028	96,222	0,585
346	papa	0,00389	0,056	96,278	1,170
347	obispo	0,00389	0,056	96,333	1,170
348	asesor fiscal	0,00380	0,028	96,361	0,585
349	conductor de coches de caballos	0,00380	0,028	96,389	0,585
350	frigorista	0,00380	0,028	96,417	0,585
351	mentor	0,00380	0,028	96,444	0,585
352	mozo de escuadra	0,00380	0,028	96,472	0,585
353	muletero	0,00380	0,028	96,500	0,585
354	recolector de frutos	0,00380	0,028	96,528	0,585
355	administrador de empresas	0,00362	0,028	96,556	0,585
356	banderillero	0,00362	0,028	96,583	0,585
357	barrista	0,00362	0,028	96,611	0,585
358	comunicador	0,00362	0,028	96,639	0,585
359	ertzaintza	0,00362	0,028	96,667	0,585
360	oftalmólogo	0,00362	0,028	96,694	0,585
361	redactor	0,00362	0,028	96,722	0,585
362	vidente	0,00362	0,028	96,750	0,585
363	vulcanólogo	0,00362	0,028	96,778	0,585
364	cardenal	0,00362	0,056	96,833	1,170
365	almacenista	0,00345	0,028	96,861	0,585
366	apoderado	0,00345	0,028	96,889	0,585
367	bedel	0,00345	0,028	96,917	0,585
368	buzo	0,00345	0,028	96,944	0,585
369	cobrador	0,00345	0,028	96,972	0,585

370	decano	0,00345	0,028	97,000	0,585
371	desempleado	0,00345	0,028	97,028	0,585
372	ebanista	0,00345	0,028	97,056	0,585
373	equilibrista	0,00345	0,028	97,083	0,585
374	médium	0,00345	0,028	97,111	0,585
375	subdirector	0,00345	0,028	97,139	0,585
376	comentarista	0,00329	0,028	97,167	0,585
377	geólogo	0,00329	0,028	97,194	0,585
378	jefe de personal	0,00329	0,028	97,222	0,585
379	recadero	0,00329	0,028	97,250	0,585
380	zapador	0,00329	0,028	97,278	0,585
381	agente	0,00314	0,028	97,306	0,585
382	agente de seguros	0,00314	0,028	97,333	0,585
383	alicatador	0,00314	0,028	97,361	0,585
384	ayudante de mecánico	0,00314	0,028	97,389	0,585
385	concertista	0,00314	0,028	97,417	0,585
386	mecanógrafo	0,00314	0,028	97,444	0,585
387	relaciones públicas	0,00314	0,028	97,472	0,585
388	tatuador	0,00314	0,028	97,500	0,585
389	vigilante de la playa	0,00314	0,028	97,528	0,585
390	alergólogo	0,00299	0,028	97,556	0,585
391	animador social	0,00299	0,028	97,583	0,585
392	antenista	0,00299	0,028	97,611	0,585
393	aprendiz de mecánico	0,00299	0,028	97,639	0,585
394	diseñador de moda	0,00299	0,028	97,667	0,585
395	escenógrafo	0,00299	0,028	97,694	0,585
396	estafador	0,00299	0,028	97,722	0,585
397	jugador de deporte	0,00299	0,028	97,750	0,585
398	marroquiner	0,00299	0,028	97,778	0,585
399	montador	0,00299	0,028	97,806	0,585
400	probador de productos	0,00299	0,028	97,833	0,585
401	solador	0,00299	0,028	97,861	0,585
402	solista	0,00299	0,028	97,889	0,585
403	terapeuta	0,00299	0,028	97,917	0,585
404	aeronauta	0,00285	0,028	97,944	0,585
405	cazador	0,00285	0,028	97,972	0,585
406	disc-jockey	0,00285	0,028	98,000	0,585
407	maquetador	0,00285	0,028	98,028	0,585
408	pirata informático	0,00285	0,028	98,056	0,585
409	tenor	0,00285	0,028	98,083	0,585
410	cuidador	0,00272	0,028	98,111	0,585
411	deportista de élite	0,00272	0,028	98,139	0,585
412	despido	0,00272	0,028	98,167	0,585
413	director de finanzas	0,00272	0,028	98,194	0,585

414	ferrallista	0,00272	0,028	98,222	0,585
415	guarda	0,00272	0,028	98,250	0,585
416	humorista	0,00272	0,028	98,278	0,585
417	animador deportivo	0,00259	0,028	98,306	0,585
418	bailador	0,00259	0,028	98,333	0,585
419	contrato	0,00259	0,028	98,361	0,585
420	controlador	0,00259	0,028	98,389	0,585
421	dramaturgo	0,00259	0,028	98,417	0,585
422	empleado de hogar	0,00259	0,028	98,444	0,585
423	especulador financiero	0,00259	0,028	98,472	0,585
424	golfista	0,00259	0,028	98,500	0,585
425	jefe de cocina	0,00259	0,028	98,528	0,585
426	nadador	0,00259	0,028	98,556	0,585
427	promotor	0,00259	0,028	98,583	0,585
428	quiromasajista	0,00259	0,028	98,611	0,585
429	taquillero	0,00259	0,028	98,639	0,585
430	técnico de rayos	0,00259	0,028	98,667	0,585
431	teniente alcalde	0,00259	0,028	98,694	0,585
432	yesista	0,00259	0,028	98,722	0,585
433	alistador	0,00247	0,028	98,750	0,585
434	asesor financiero	0,00247	0,028	98,778	0,585
435	chófer de camión	0,00247	0,028	98,806	0,585
436	escaparatista	0,00247	0,028	98,833	0,585
437	estríper	0,00247	0,028	98,861	0,585
438	folclórico	0,00247	0,028	98,889	0,585
439	gobernador	0,00247	0,028	98,917	0,585
440	huelga	0,00247	0,028	98,944	0,585
441	jefe de estación	0,00247	0,028	98,972	0,585
442	paisajista	0,00247	0,028	99,000	0,585
443	porquero	0,00247	0,028	99,028	0,585
444	trovador	0,00247	0,028	99,056	0,585
445	aprendiz	0,00235	0,028	99,083	0,585
446	artesano	0,00235	0,028	99,111	0,585
447	barman	0,00235	0,028	99,139	0,585
448	buceador	0,00235	0,028	99,167	0,585
449	cineasta	0,00235	0,028	99,194	0,585
450	embajador	0,00235	0,028	99,222	0,585
451	maestro de obras	0,00235	0,028	99,250	0,585
452	patrón	0,00235	0,028	99,278	0,585
453	psicoanalista	0,00235	0,028	99,306	0,585
454	sargento	0,00235	0,028	99,333	0,585
455	vaquero	0,00235	0,028	99,361	0,585
456	alfarero	0,00224	0,028	99,389	0,585
457	fraile	0,00224	0,028	99,417	0,585
458	personal doméstico	0,00224	0,028	99,444	0,585

459	profesor de yoga	0,00224	0,028	99,472	0,585
460	proxeneta	0,00224	0,028	99,500	0,585
461	técnico de imagen y sonido	0,00224	0,028	99,528	0,585
462	criado	0,00214	0,028	99,556	0,585
463	fabricante	0,00214	0,028	99,583	0,585
464	maestro de reiki	0,00214	0,028	99,611	0,585
465	monaguillo	0,00214	0,028	99,639	0,585
466	armero	0,00204	0,028	99,667	0,585
467	boxeador	0,00204	0,028	99,694	0,585
468	monje	0,00204	0,028	99,722	0,585
469	azafato de vuelos	0,00194	0,028	99,750	0,585
470	coordinador	0,00194	0,028	99,778	0,585
471	mozo	0,00194	0,028	99,806	0,585
472	oficial de la marina	0,00194	0,028	99,833	0,585
473	delegado	0,00185	0,028	99,861	0,585
474	oficial	0,00185	0,028	99,889	0,585
475	reportero	0,00185	0,028	99,917	0,585
476	zoólogo	0,00177	0,028	99,944	0,585
477	inmobiliario	0,00153	0,028	99,972	0,585
478	rector	0,00146	0,028	100,000	0,585

1.2.9. Ordenadores e internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	ratón	0,68471	4,067	4,067	77,907
2	teclado	0,66352	4,097	8,164	78,488
3	pantalla	0,61087	3,733	11,897	71,512
4	página web	0,28249	2,124	14,021	40,698
5	portátil	0,26578	1,760	15,781	33,721
6	correo (electrónico)	0,24676	1,973	17,754	37,791
7	disco duro	0,24014	1,942	19,697	37,209
8	torre	0,22941	1,700	21,396	32,558
9	red social	0,21344	1,730	23,126	33,140
10	Facebook	0,20834	1,608	24,734	30,814
11	programa	0,20668	1,700	26,434	32,558
12	impresora	0,20628	1,548	27,982	29,651
13	monitor	0,19647	1,335	29,317	25,581
14	cable	0,17350	1,366	30,683	26,163
15	archivo	0,16377	1,305	31,988	25,000
16	software	0,16137	1,305	33,293	25,000
17	pendrive	0,15910	1,275	34,568	24,419
18	hardware	0,15544	1,244	35,812	23,837
19	Google	0,14110	1,184	36,995	22,674
20	Tuenti	0,13948	1,062	38,058	20,349
21	buscador	0,13783	1,153	39,211	22,093
22	escáner	0,13177	1,093	40,303	20,930
23	chat	0,12461	1,062	41,366	20,349
24	USB	0,11299	0,910	42,276	17,442
25	blog	0,11262	0,880	43,156	16,860
26	CD(-ROM)	0,09981	0,850	44,006	16,279
27	virus	0,09896	0,880	44,886	16,860
28	wifi	0,09710	0,850	45,736	16,279
29	ordenador	0,09654	0,546	46,282	10,465
30	conexión	0,09108	0,789	47,071	15,116
31	altavoz	0,09101	0,698	47,769	13,372
32	memoria	0,08892	0,698	48,467	13,372
33	módem	0,08891	0,789	49,256	15,116
34	router	0,08640	0,759	50,015	14,535
35	carpetas	0,08567	0,728	50,744	13,953
36	Windows	0,08399	0,637	51,381	12,209
37	red	0,08228	0,637	52,018	12,209
38	internet	0,08004	0,577	52,595	11,047
39	chatear	0,07561	0,577	53,171	11,047
40	descargar	0,07546	0,607	53,778	11,628
41	antivirus	0,07319	0,698	54,476	13,372
42	placa base	0,07313	0,577	55,053	11,047
43	navegar	0,07005	0,546	55,599	10,465
44	servidor	0,06827	0,607	56,206	11,628
45	tecla	0,06655	0,486	56,692	9,302

46	web	0,06627	0,486	57,178	9,302
47	descarga	0,06511	0,577	57,754	11,047
48	DVD	0,06405	0,546	58,300	10,465
49	foro	0,06165	0,516	58,816	9,884
50	Twitter	0,06047	0,455	59,272	8,721
51	documento	0,05840	0,516	59,788	9,884
52	Hotmail	0,05771	0,546	60,334	10,465
53	batería	0,05597	0,455	60,789	8,721
54	tarjeta gráfica	0,05560	0,486	61,275	9,302
55	(Microsoft) Word	0,05469	0,455	61,730	8,721
56	memoria RAM	0,05349	0,516	62,246	9,884
57	procesador de textos	0,05344	0,455	62,701	8,721
58	e-mail	0,05202	0,425	63,126	8,140
59	cortafuego	0,05172	0,455	63,581	8,721
60	puerto USB	0,05000	0,425	64,006	8,140
61	Wikipedia	0,04926	0,425	64,431	8,140
62	CPU	0,04878	0,364	64,795	6,977
63	escritorio	0,04821	0,425	65,220	8,140
64	Youtube	0,04771	0,395	65,615	7,558
65	grabadora	0,04770	0,395	66,009	7,558
66	PC	0,04703	0,273	66,282	5,233
67	sistema operativo	0,04577	0,364	66,646	6,977
68	campus virtual	0,04384	0,303	66,950	5,814
69	Messenger	0,04280	0,364	67,314	6,977
70	navegador	0,04230	0,364	67,678	6,977
71	formatear	0,04203	0,334	68,012	6,395
72	alfombrilla	0,04138	0,303	68,316	5,814
73	ventana	0,03775	0,303	68,619	5,814
74	ADSL	0,03723	0,303	68,923	5,814
75	contraseña	0,03636	0,273	69,196	5,233
76	wiki	0,03627	0,303	69,499	5,814
77	cargador	0,03589	0,273	69,772	5,233
78	disco	0,03455	0,273	70,046	5,233
79	disquete	0,03389	0,273	70,319	5,233
80	enlace	0,03353	0,303	70,622	5,814
81	ordenador de sobremesa	0,03314	0,212	70,835	4,070
82	enchufe	0,03287	0,273	71,108	5,233
83	juego	0,03272	0,303	71,411	5,814
84	información	0,03225	0,273	71,684	5,233
85	cámara web	0,03160	0,303	71,988	5,814
86	(Microsoft) PowerPoint	0,03120	0,273	72,261	5,233
87	Linux	0,03095	0,243	72,504	4,651
88	Hot Potatoes	0,03059	0,273	72,777	5,233
89	videojuego	0,03013	0,212	72,989	4,070
90	música	0,02939	0,243	73,232	4,651
91	vídeo	0,02779	0,243	73,475	4,651

92	inalámbrico	0,02758	0,273	73,748	5,233
93	descargar archivo	0,02755	0,182	73,930	3,488
94	sitio web	0,02732	0,212	74,143	4,070
95	subir	0,02726	0,243	74,385	4,651
96	instalar	0,02645	0,243	74,628	4,651
97	Microsoft	0,02619	0,212	74,841	4,070
98	procesador	0,02463	0,243	75,083	4,651
99	netbook	0,02446	0,152	75,235	2,907
100	(Microsoft) Excel	0,02436	0,212	75,448	4,070
101	Yahoo	0,02428	0,243	75,690	4,651
102	cookie	0,02416	0,243	75,933	4,651
103	red inalámbrica	0,02386	0,212	76,146	4,070
104	página	0,02362	0,182	76,328	3,488
105	TIC	0,02348	0,152	76,480	2,907
106	película	0,02236	0,212	76,692	4,070
107	videocámara	0,02231	0,182	76,874	3,488
108	Macintosh	0,02218	0,182	77,056	3,488
109	base de datos	0,02190	0,182	77,238	3,488
110	Apple	0,01995	0,152	77,390	2,907
111	JClic	0,01993	0,182	77,572	3,488
112	salvapantalla	0,01992	0,152	77,724	2,907
113	banda ancha	0,01834	0,121	77,845	2,326
114	comunicación	0,01808	0,152	77,997	2,907
115	portal	0,01808	0,152	78,149	2,907
116	(Microsoft) Internet Explorer	0,01763	0,152	78,300	2,907
117	webcam	0,01752	0,182	78,483	3,488
118	microprocesador	0,01739	0,182	78,665	3,488
119	fichero	0,01736	0,121	78,786	2,326
120	eMule	0,01665	0,182	78,968	3,488
121	computadora	0,01586	0,091	79,059	1,744
122	lápiz de memoria	0,01572	0,152	79,211	2,907
123	notebook	0,01557	0,121	79,332	2,326
124	bajar archivo	0,01542	0,121	79,454	2,326
125	surfear	0,01533	0,121	79,575	2,326
126	URL	0,01529	0,121	79,697	2,326
127	tarjeta de memoria	0,01505	0,121	79,818	2,326
128	compilador	0,01475	0,121	79,939	2,326
129	bajar	0,01463	0,091	80,030	1,744
130	usuario	0,01450	0,152	80,182	2,907
131	aplicación	0,01449	0,152	80,334	2,907
132	guardar	0,01440	0,121	80,455	2,326
133	pdf	0,01432	0,152	80,607	2,907
134	copiar	0,01427	0,152	80,759	2,907
135	velocidad	0,01421	0,121	80,880	2,326
136	clic	0,01407	0,121	81,002	2,326
137	IP	0,01396	0,121	81,123	2,326
138	AutoCAD	0,01388	0,121	81,244	2,326
139	Gmail	0,01360	0,121	81,366	2,326

140	piratear	0,01358	0,121	81,487	2,326
141	mensaje	0,01349	0,121	81,608	2,326
142	Toshiba	0,01347	0,091	81,700	1,744
143	icono	0,01344	0,152	81,851	2,907
144	Java	0,01322	0,091	81,942	1,744
145	ventilador	0,01313	0,152	82,094	2,907
146	mega	0,01303	0,121	82,215	2,326
147	programar	0,01288	0,152	82,367	2,907
148	MSN	0,01243	0,121	82,489	2,326
149	tarjeta de sonido	0,01233	0,091	82,580	1,744
150	programa informático	0,01232	0,091	82,671	1,744
151	fuelle de alimentación	0,01231	0,121	82,792	2,326
152	megabyte	0,01199	0,121	82,914	2,326
153	imprimir	0,01178	0,091	83,005	1,744
154	Moodle	0,01175	0,091	83,096	1,744
155	ordenador personal	0,01163	0,061	83,156	1,163
156	directorio	0,01160	0,121	83,278	2,326
157	casco	0,01155	0,091	83,369	1,744
158	Skype	0,01154	0,091	83,460	1,744
159	disco duro externo	0,01146	0,121	83,581	2,326
160	Spotify	0,01136	0,091	83,672	1,744
161	pegar	0,01107	0,121	83,794	2,326
162	protocolo	0,01097	0,121	83,915	2,326
163	buscar	0,01097	0,091	84,006	1,744
164	desfragmentar	0,01086	0,091	84,097	1,744
165	cargar	0,01085	0,061	84,158	1,163
166	Series Yonkis	0,01024	0,091	84,249	1,744
167	byte	0,01023	0,091	84,340	1,744
168	cámara	0,01017	0,091	84,431	1,744
169	clave	0,01014	0,091	84,522	1,744
170	dirección	0,01010	0,121	84,643	2,326
171	RAM	0,01008	0,061	84,704	1,163
172	periférico	0,00999	0,091	84,795	1,744
173	dirección web	0,00982	0,091	84,886	1,744
174	conector	0,00982	0,091	84,977	1,744
175	grabar	0,00960	0,091	85,068	1,744
176	ejecutar	0,00956	0,091	85,159	1,744
177	intranet	0,00944	0,091	85,250	1,744
178	barra de herramientas	0,00938	0,091	85,341	1,744
179	subir archivos	0,00927	0,091	85,432	1,744
180	iPad	0,00922	0,091	85,524	1,744
181	bandeja de entrada	0,00915	0,121	85,645	2,326
182	memoria interna	0,00873	0,061	85,706	1,163
183	historial	0,00867	0,091	85,797	1,744
184	lápiz óptico	0,00855	0,061	85,857	1,163
185	programador	0,00853	0,091	85,948	1,744

186	Dell	0,00852	0,061	86,009	1,163
187	lenguaje informático	0,00840	0,061	86,070	1,163
188	controlador	0,00840	0,061	86,131	1,163
189	escanear	0,00832	0,061	86,191	1,163
190	buscar información	0,00831	0,061	86,252	1,163
191	técnico	0,00816	0,091	86,343	1,744
192	juego online	0,00815	0,061	86,404	1,163
193	proyector	0,00814	0,061	86,464	1,163
194	adjuntar	0,00813	0,091	86,555	1,744
195	memoria externa	0,00807	0,061	86,616	1,163
196	WebQuest	0,00807	0,061	86,677	1,163
197	Paint	0,00803	0,091	86,768	1,744
198	unidad central	0,00796	0,061	86,829	1,163
199	tarjeta de red	0,00785	0,061	86,889	1,163
200	tinta	0,00778	0,091	86,980	1,744
201	CorelDRAW	0,00769	0,061	87,041	1,163
202	WordReference	0,00767	0,091	87,132	1,744
203	dirección de correo (electrónico)	0,00762	0,061	87,193	1,163
204	gigabyte	0,00758	0,091	87,284	1,744
205	pirata informático	0,00756	0,061	87,344	1,163
206	placa madre	0,00755	0,061	87,405	1,163
207	configurar	0,00749	0,091	87,496	1,744
208	noticia	0,00748	0,091	87,587	1,744
209	(Microsoft) Office	0,00725	0,061	87,648	1,163
210	tableta digitalizadora	0,00721	0,061	87,709	1,163
211	(Microsoft) Access	0,00708	0,061	87,769	1,163
212	streaming	0,00708	0,061	87,830	1,163
213	chip	0,00693	0,061	87,891	1,163
214	videoconferencia	0,00687	0,061	87,951	1,163
215	botón	0,00674	0,091	88,042	1,744
216	tarjeta de vídeo	0,00670	0,061	88,103	1,163
217	cortar	0,00656	0,061	88,164	1,163
218	ratón inalámbrico	0,00654	0,061	88,225	1,163
219	hoja de cálculo	0,00654	0,061	88,285	1,163
220	imagen	0,00642	0,061	88,346	1,163
221	Google Maps	0,00639	0,061	88,407	1,163
222	www	0,00628	0,061	88,467	1,163
223	desinstalar	0,00625	0,061	88,528	1,163
224	caja	0,00624	0,061	88,589	1,163
225	página principal	0,00613	0,061	88,649	1,163
226	transistor	0,00613	0,061	88,710	1,163
227	herramienta	0,00612	0,061	88,771	1,163
228	pinchar	0,00609	0,061	88,832	1,163
229	palanca de mando	0,00609	0,061	88,892	1,163
230	mensajería	0,00605	0,061	88,953	1,163
231	papelera de reciclaje	0,00601	0,061	89,014	1,163
232	Safari	0,00600	0,061	89,074	1,163
233	almacenamiento	0,00583	0,061	89,135	1,163

234	código	0,00581	0,030	89,165	0,581
235	freaky	0,00581	0,030	89,196	0,581
236	MS-DOS	0,00581	0,030	89,226	0,581
237	aula virtual	0,00581	0,030	89,256	0,581
238	proceso	0,00573	0,061	89,317	1,163
239	bluetooth	0,00564	0,061	89,378	1,163
240	compilar	0,00562	0,061	89,439	1,163
241	El Tiempo	0,00562	0,061	89,499	1,163
242	Ajax	0,00554	0,030	89,530	0,581
243	Europa FM	0,00554	0,030	89,560	0,581
244	nombre de usuario	0,00554	0,030	89,590	0,581
245	interfaz	0,00550	0,061	89,651	1,163
246	puntero	0,00546	0,061	89,712	1,163
247	condensador	0,00541	0,061	89,772	1,163
248	disquetera	0,00535	0,061	89,833	1,163
249	Acer	0,00528	0,030	89,863	0,581
250	corriente	0,00528	0,030	89,894	0,581
251	Gran Hermano	0,00528	0,030	89,924	0,581
252	Intel Core	0,00528	0,030	89,954	0,581
253	lenguaje de programación	0,00528	0,030	89,985	0,581
254	videollamada	0,00528	0,030	90,015	0,581
255	cliente	0,00527	0,061	90,076	1,163
256	DNS	0,00524	0,061	90,137	1,163
257	Adobe Photoshop	0,00519	0,061	90,197	1,163
258	informática	0,00519	0,061	90,258	1,163
259	color	0,00514	0,061	90,319	1,163
260	Mozilla Firefox	0,00514	0,061	90,379	1,163
261	C++	0,00504	0,030	90,410	0,581
262	descargar documento	0,00504	0,030	90,440	0,581
263	lápiz USB	0,00504	0,030	90,470	0,581
264	marcador	0,00503	0,061	90,531	1,163
265	tarjeta	0,00496	0,061	90,592	1,163
266	formato jpg	0,00491	0,061	90,653	1,163
267	resolución	0,00490	0,061	90,713	1,163
268	teléfono	0,00490	0,061	90,774	1,163
269	ofimática	0,00487	0,061	90,835	1,163
270	administrador de tareas	0,00480	0,030	90,865	0,581
271	conectividad	0,00480	0,030	90,895	0,581
272	cursor	0,00480	0,030	90,926	0,581
273	Farm Ville	0,00480	0,030	90,956	0,581
274	fax	0,00480	0,030	90,986	0,581
275	función	0,00480	0,030	91,017	0,581
276	HTML	0,00480	0,030	91,047	0,581
277	intérprete de	0,00480	0,030	91,077	0,581

	comandos				
278	LinkedIn	0,00480	0,030	91,108	0,581
279	micrófono	0,00480	0,030	91,138	0,581
280	página oficial	0,00480	0,030	91,168	0,581
281	tarjeta de audio	0,00480	0,030	91,199	0,581
282	tensión	0,00480	0,030	91,229	0,581
283	Uma	0,00480	0,030	91,259	0,581
284	UNIX	0,00480	0,030	91,290	0,581
285	comando	0,00473	0,061	91,351	1,163
286	insertar	0,00472	0,061	91,411	1,163
287	sistema	0,00468	0,061	91,472	1,163
288	diseño gráfico	0,00464	0,061	91,533	1,163
289	campus	0,00458	0,030	91,563	0,581
290	electricidad	0,00458	0,030	91,593	0,581
291	enciclopedia virtual	0,00458	0,030	91,624	0,581
292	FriendFeed	0,00458	0,030	91,654	0,581
293	GNU	0,00458	0,030	91,684	0,581
294	InfoJobs	0,00458	0,030	91,715	0,581
295	letra	0,00458	0,030	91,745	0,581
296	librería	0,00458	0,030	91,775	0,581
297	página web de contenido erótico	0,00458	0,030	91,806	0,581
298	periódico	0,00458	0,030	91,836	0,581
299	raíz	0,00458	0,030	91,866	0,581
300	Sony Vaio	0,00458	0,030	91,897	0,581
301	troyano	0,00458	0,030	91,927	0,581
302	instrucción	0,00451	0,061	91,988	1,163
303	BitTorrent	0,00446	0,061	92,049	1,163
304	puerto	0,00446	0,061	92,109	1,163
305	Ares	0,00437	0,061	92,170	1,163
306	conectarse	0,00436	0,030	92,200	0,581
307	entretenimiento	0,00436	0,030	92,231	0,581
308	extraíble	0,00436	0,030	92,261	0,581
309	global	0,00436	0,030	92,291	0,581
310	Google Earth	0,00436	0,030	92,322	0,581
311	Google Reader	0,00436	0,030	92,352	0,581
312	grupo internauta	0,00436	0,030	92,382	0,581
313	informático	0,00436	0,030	92,413	0,581
314	LibreOffice	0,00436	0,030	92,443	0,581
315	objeto	0,00436	0,030	92,473	0,581
316	simulador de circuitos	0,00436	0,030	92,504	0,581
317	sonido	0,00436	0,030	92,534	0,581
318	transformador	0,00436	0,030	92,564	0,581
319	actualizar	0,00430	0,061	92,625	1,163
320	megapíxel	0,00430	0,061	92,686	1,163

321	trabajo	0,00430	0,061	92,747	1,163
322	bajar música	0,00416	0,030	92,777	0,581
323	casete	0,00416	0,030	92,807	0,581
324	conexión telefónica	0,00416	0,030	92,838	0,581
325	configuración	0,00416	0,030	92,868	0,581
326	dato	0,00416	0,030	92,898	0,581
327	disco regrabable	0,00416	0,030	92,929	0,581
328	HP	0,00416	0,030	92,959	0,581
329	mando a distancia	0,00416	0,030	92,989	0,581
330	manual	0,00416	0,030	93,020	0,581
331	Pentium	0,00416	0,030	93,050	0,581
332	Proxy	0,00416	0,030	93,080	0,581
333	terminal virtual	0,00416	0,030	93,111	0,581
334	tóner	0,00416	0,030	93,141	0,581
335	unidad de control	0,00416	0,030	93,171	0,581
336	World Wide Web	0,00416	0,030	93,202	0,581
337	ancho de banda	0,00396	0,030	93,232	0,581
338	auricular	0,00396	0,030	93,263	0,581
339	cable de serie	0,00396	0,030	93,293	0,581
340	cartucho de tóner	0,00396	0,030	93,323	0,581
341	comprimir	0,00396	0,030	93,354	0,581
342	contacto	0,00396	0,030	93,384	0,581
343	copiar documento	0,00396	0,030	93,414	0,581
344	inicio	0,00396	0,030	93,445	0,581
345	papelera	0,00396	0,030	93,475	0,581
346	RIP	0,00396	0,030	93,505	0,581
347	rutina	0,00396	0,030	93,536	0,581
348	SlideShare	0,00396	0,030	93,566	0,581
349	tarjeta sintonizadora de televisión	0,00396	0,030	93,596	0,581
350	teclado numérico	0,00396	0,030	93,627	0,581
351	vínculo	0,00396	0,030	93,657	0,581
352	spam	0,00396	0,061	93,718	1,163
353	favorito	0,00396	0,061	93,778	1,163
354	BIN	0,00378	0,030	93,809	0,581
355	buscar en la red	0,00378	0,030	93,839	0,581
356	búsqueda	0,00378	0,030	93,869	0,581
357	descomprimir	0,00378	0,030	93,900	0,581
358	disco compacto	0,00378	0,030	93,930	0,581
359	guardar documento	0,00378	0,030	93,961	0,581
360	hacker	0,00378	0,030	93,991	0,581
361	libro electrónico	0,00378	0,030	94,021	0,581
362	memoria USB	0,00378	0,030	94,052	0,581
363	programa operativo	0,00378	0,030	94,082	0,581
364	proveedor	0,00378	0,030	94,112	0,581
365	red de alta velocidad	0,00378	0,030	94,143	0,581

366	Audacity	0,00360	0,030	94,173	0,581
367	barra	0,00360	0,030	94,203	0,581
368	conmutador	0,00360	0,030	94,234	0,581
369	distracción	0,00360	0,030	94,264	0,581
370	documento de texto	0,00360	0,030	94,294	0,581
371	Dropbox	0,00360	0,030	94,325	0,581
372	escribir	0,00360	0,030	94,355	0,581
373	fondo de pantalla	0,00360	0,030	94,385	0,581
374	iPhone	0,00360	0,030	94,416	0,581
375	mecanografía	0,00360	0,030	94,446	0,581
376	Mis documentos	0,00360	0,030	94,476	0,581
377	pestaña	0,00360	0,030	94,507	0,581
378	Protocolo HTTP	0,00360	0,030	94,537	0,581
379	red 2.0	0,00360	0,030	94,568	0,581
380	reportaje	0,00360	0,030	94,598	0,581
381	tableta gráfica	0,00360	0,030	94,628	0,581
382	Mi PC	0,00351	0,061	94,689	1,163
383	seguridad	0,00351	0,061	94,750	1,163
384	adaptador inalámbrico	0,00343	0,030	94,780	0,581
385	Android	0,00343	0,030	94,810	0,581
386	antena	0,00343	0,030	94,841	0,581
387	borrar	0,00343	0,030	94,871	0,581
388	calendario	0,00343	0,030	94,901	0,581
389	control parental	0,00343	0,030	94,932	0,581
390	descompilar	0,00343	0,030	94,962	0,581
391	disco extraíble	0,00343	0,030	94,992	0,581
392	ejecutable	0,00343	0,030	95,023	0,581
393	gateway	0,00343	0,030	95,053	0,581
394	Guadalinex	0,00343	0,030	95,083	0,581
395	gusano	0,00343	0,030	95,114	0,581
396	MP3	0,00343	0,030	95,144	0,581
397	multifunción	0,00343	0,030	95,175	0,581
398	instalación	0,00334	0,061	95,235	1,163
399	audio	0,00327	0,030	95,266	0,581
400	consola	0,00327	0,030	95,296	0,581
401	copiadora	0,00327	0,030	95,326	0,581
402	extensión	0,00327	0,030	95,357	0,581
403	mesa	0,00327	0,030	95,387	0,581
404	nuevas tecnologías	0,00327	0,030	95,417	0,581
405	píxel	0,00327	0,030	95,448	0,581
406	resistencia	0,00327	0,030	95,478	0,581
407	sitio	0,00327	0,030	95,508	0,581
408	solitario	0,00327	0,030	95,539	0,581
409	visualizar	0,00327	0,030	95,569	0,581
410	Windows Movie	0,00327	0,030	95,599	0,581

	Maker				
411	árbol	0,00312	0,030	95,630	0,581
412	Avast	0,00312	0,030	95,660	0,581
413	cable de red	0,00312	0,030	95,690	0,581
414	cable USB	0,00312	0,030	95,721	0,581
415	centro TIC	0,00312	0,030	95,751	0,581
416	cine	0,00312	0,030	95,781	0,581
417	cristal de cuarzo	0,00312	0,030	95,812	0,581
418	dígito	0,00312	0,030	95,842	0,581
419	dominio web	0,00312	0,030	95,873	0,581
420	espacio virtual	0,00312	0,030	95,903	0,581
421	GPRS	0,00312	0,030	95,933	0,581
422	licencia	0,00312	0,030	95,964	0,581
423	mantenimiento	0,00312	0,030	95,994	0,581
424	medio de comunicación social	0,00312	0,030	96,024	0,581
425	página de inicio	0,00312	0,030	96,055	0,581
426	periódico virtual	0,00312	0,030	96,085	0,581
427	programación televisiva	0,00312	0,030	96,115	0,581
428	silla	0,00312	0,030	96,146	0,581
429	software libre	0,00312	0,030	96,176	0,581
430	web 2.0	0,00312	0,030	96,206	0,581
431	accesorio	0,00297	0,030	96,237	0,581
432	contenido	0,00297	0,030	96,267	0,581
433	disco portátil	0,00297	0,030	96,297	0,581
434	dominio	0,00297	0,030	96,328	0,581
435	editor de texto	0,00297	0,030	96,358	0,581
436	educación	0,00297	0,030	96,388	0,581
437	hipervínculo	0,00297	0,030	96,419	0,581
438	ingeniero	0,00297	0,030	96,449	0,581
439	inscripción	0,00297	0,030	96,480	0,581
440	Intro	0,00297	0,030	96,510	0,581
441	línea	0,00297	0,030	96,540	0,581
442	operador de búsqueda	0,00297	0,030	96,571	0,581
443	Picassa	0,00297	0,030	96,601	0,581
444	serie	0,00297	0,030	96,631	0,581
445	sistema binario	0,00297	0,030	96,662	0,581
446	suscripción	0,00297	0,030	96,692	0,581
447	videojuego online	0,00297	0,030	96,722	0,581
448	casilla	0,00283	0,030	96,753	0,581
449	diapositiva	0,00283	0,030	96,783	0,581
450	electrónica	0,00283	0,030	96,813	0,581
451	Google Crome	0,00283	0,030	96,844	0,581
452	hackear	0,00283	0,030	96,874	0,581
453	Hofmann	0,00283	0,030	96,904	0,581

454	lector de CD-ROM	0,00283	0,030	96,935	0,581
455	parque tecnológico	0,00283	0,030	96,965	0,581
456	podcast	0,00283	0,030	96,995	0,581
457	retardo	0,00283	0,030	97,026	0,581
458	seleccionar	0,00283	0,030	97,056	0,581
459	vídeo musical	0,00283	0,030	97,086	0,581
460	adaptador	0,00270	0,030	97,117	0,581
461	conocer gente	0,00270	0,030	97,147	0,581
462	diccionario	0,00270	0,030	97,178	0,581
463	FireWire	0,00270	0,030	97,208	0,581
464	formato de negrita	0,00270	0,030	97,238	0,581
465	FTP	0,00270	0,030	97,269	0,581
466	hacer clic	0,00270	0,030	97,299	0,581
467	panel de control	0,00270	0,030	97,329	0,581
468	puesto	0,00270	0,030	97,360	0,581
469	Rincón del Vago	0,00270	0,030	97,390	0,581
470	RSS	0,00270	0,030	97,420	0,581
471	script	0,00270	0,030	97,451	0,581
472	unidad de disco	0,00270	0,030	97,481	0,581
473	botón derecho	0,00257	0,030	97,511	0,581
474	conector DIN	0,00257	0,030	97,542	0,581
475	disco duro externo				
	multimedia	0,00257	0,030	97,572	0,581
476	enciclopedia	0,00257	0,030	97,602	0,581
477	estilo	0,00257	0,030	97,633	0,581
478	hibernar	0,00257	0,030	97,663	0,581
479	Interplex	0,00257	0,030	97,693	0,581
480	output	0,00257	0,030	97,724	0,581
481	portafolio	0,00257	0,030	97,754	0,581
482	tableta	0,00257	0,030	97,785	0,581
483	aceptar	0,00245	0,030	97,815	0,581
484	ayuda	0,00245	0,030	97,845	0,581
485	conexión ADSL	0,00245	0,030	97,876	0,581
486	documental	0,00245	0,030	97,906	0,581
487	ensamblador	0,00245	0,030	97,936	0,581
488	Ethernet	0,00245	0,030	97,967	0,581
489	formato	0,00245	0,030	97,997	0,581
490	online	0,00245	0,030	98,027	0,581
491	SDL Trados	0,00245	0,030	98,058	0,581
492	tarifa plana	0,00245	0,030	98,088	0,581
493	unidad de CD	0,00245	0,030	98,118	0,581
494	bit	0,00234	0,030	98,149	0,581
495	consulta	0,00234	0,030	98,179	0,581
496	fondo	0,00234	0,030	98,209	0,581
497	grabador de CD	0,00234	0,030	98,240	0,581
498	interactivo	0,00234	0,030	98,270	0,581

499	lápiz táctil	0,00234	0,030	98,300	0,581
500	Sin la Mula	0,00234	0,030	98,331	0,581
501	teleformación	0,00234	0,030	98,361	0,581
502	Terra	0,00234	0,030	98,392	0,581
503	acceso	0,00223	0,030	98,422	0,581
504	amplitud	0,00223	0,030	98,452	0,581
505	dispositivo	0,00223	0,030	98,483	0,581
506	palabra	0,00223	0,030	98,513	0,581
507	servidor web	0,00223	0,030	98,543	0,581
508	teletrabajo	0,00223	0,030	98,574	0,581
509	televisión	0,00223	0,030	98,604	0,581
510	Vodafone	0,00223	0,030	98,634	0,581
511	clavija	0,00213	0,030	98,665	0,581
512	firewall	0,00213	0,030	98,695	0,581
513	interminable	0,00213	0,030	98,725	0,581
514	iPod	0,00213	0,030	98,756	0,581
515	máscara	0,00213	0,030	98,786	0,581
516	Orange	0,00213	0,030	98,816	0,581
517	Telefónica	0,00213	0,030	98,847	0,581
518	virtual	0,00213	0,030	98,877	0,581
519	antiespía	0,00203	0,030	98,907	0,581
520	control remoto	0,00203	0,030	98,938	0,581
521	enviar	0,00203	0,030	98,968	0,581
522	interrupción	0,00203	0,030	98,998	0,581
523	multimedia	0,00203	0,030	99,029	0,581
524	perfil	0,00203	0,030	99,059	0,581
525	portal web	0,00203	0,030	99,090	0,581
526	publicidad	0,00203	0,030	99,120	0,581
527	social	0,00203	0,030	99,150	0,581
528	TCP	0,00203	0,030	99,181	0,581
529	agregar contacto	0,00193	0,030	99,211	0,581
530	cultura	0,00193	0,030	99,241	0,581
531	inventor	0,00193	0,030	99,272	0,581
532	PDA	0,00193	0,030	99,302	0,581
533	ranura	0,00193	0,030	99,332	0,581
534	administrador	0,00184	0,030	99,363	0,581
535	amistad	0,00184	0,030	99,393	0,581
536	burbuja	0,00184	0,030	99,423	0,581
537	cibernauta	0,00184	0,030	99,454	0,581
538	error	0,00184	0,030	99,484	0,581
539	lengua	0,00184	0,030	99,514	0,581
540	depurar	0,00175	0,030	99,545	0,581
541	intercambio	0,00175	0,030	99,575	0,581
542	paralelismo	0,00175	0,030	99,605	0,581
543	problema	0,00167	0,030	99,636	0,581
544	rapidez	0,00167	0,030	99,666	0,581

545	serie online	0,00167	0,030	99,697	0,581
546	arreglo	0,00159	0,030	99,727	0,581
547	descarga directa	0,00159	0,030	99,757	0,581
548	terabyte	0,00159	0,030	99,788	0,581
549	acceder	0,00145	0,030	99,818	0,581
550	codificación	0,00145	0,030	99,848	0,581
551	doble clic	0,00145	0,030	99,879	0,581
552	prensa	0,00145	0,030	99,909	0,581
553	HDMI	0,00132	0,030	99,939	0,581
554	certificado	0,00120	0,030	99,970	0,581
555	Windows Vista	0,00114	0,030	100,000	0,581

2. Diccionarios en inglés

2.1. Listado de palabras en inglés por orden alfabético

2.1.1. Human body parts

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(eye)lash	0,07289	0,856	0,856	14,667
2	abdomen	0,00281	0,039	0,895	0,667
3	ankle	0,07161	0,817	1,711	14,000
4	arm	0,57588	4,706	6,418	80,667
5	armpit	0,00508	0,078	6,496	1,333
6	artery	0,00262	0,039	6,534	0,667
7	ass	0,05433	0,700	7,235	12,000
8	axis	0,00183	0,039	7,273	0,667
9	back	0,13911	1,906	9,179	32,667
10	ball	0,00183	0,039	9,218	0,667
11	belly	0,03617	0,428	9,646	7,333
12	belly button	0,01138	0,156	9,802	2,667
13	biceps	0,00450	0,078	9,879	1,333
14	blood	0,01534	0,194	10,074	3,333
15	blood vessel	0,00111	0,039	10,113	0,667
16	bob	0,00325	0,039	10,152	0,667
17	body	0,01697	0,194	10,346	3,333
18	bone	0,04860	0,739	11,085	12,667
19	bottom	0,03806	0,622	11,708	10,667
20	brain	0,04849	0,545	12,252	9,333
21	breast	0,07479	1,089	13,341	18,667
22	brow	0,00385	0,078	13,419	1,333
23	bum	0,00302	0,039	13,458	0,667
24	bump	0,00302	0,039	13,497	0,667
25	bust	0,00602	0,078	13,574	1,333
26	butt	0,00525	0,078	13,652	1,333
27	calf muscle	0,00644	0,117	13,769	2,000
28	cheek	0,05722	0,778	14,547	13,333
29	cheer	0,00500	0,039	14,586	0,667
30	chest	0,08555	1,167	15,753	20,000
31	chick	0,00227	0,039	15,792	0,667
32	chin	0,03869	0,467	16,258	8,000
33	cock	0,00227	0,039	16,297	0,667
34	column	0,00137	0,039	16,336	0,667
35	cranium	0,00325	0,039	16,375	0,667
36	cunt	0,00302	0,039	16,414	0,667
37	dick	0,00576	0,078	16,492	1,333
38	ear	0,40467	3,890	20,381	66,667
39	ear lobe	0,00735	0,117	20,498	2,000

40	elbow	0,13626	1,322	21,820	22,667
41	elbow joint	0,00147	0,039	21,859	0,667
42	epidermis	0,00349	0,039	21,898	0,667
43	eye	0,64489	5,056	26,954	86,667
44	eyebrow	0,11998	1,322	28,277	22,667
45	eyelid	0,03056	0,428	28,705	7,333
46	face	0,15769	1,439	30,144	24,667
47	feet	0,23043	2,178	32,322	37,333
48	finger	0,53468	5,056	37,378	86,667
49	fist	0,00325	0,039	37,417	0,667
50	food	0,00577	0,039	37,456	0,667
51	foot	0,25355	2,528	39,984	43,333
52	forearm	0,01640	0,194	40,179	3,333
53	forehead	0,01748	0,272	40,451	4,667
54	groin	0,00183	0,039	40,490	0,667
55	hair	0,38842	3,773	44,263	64,667
56	hamstring	0,00147	0,039	44,302	0,667
57	hand	0,52240	4,590	48,891	78,667
58	head	0,68183	4,862	53,753	83,333
59	hear	0,00229	0,078	53,831	1,333
60	heart	0,17242	2,178	56,009	37,333
61	heel	0,00465	0,039	56,048	0,667
62	hip	0,03051	0,350	56,398	6,000
63	illness	0,00128	0,039	56,437	0,667
64	intestine	0,00410	0,078	56,515	1,333
65	iris	0,00302	0,039	56,554	0,667
66	jaw	0,01394	0,194	56,748	3,333
67	joint	0,00325	0,039	56,787	0,667
68	kidney	0,03250	0,622	57,410	10,667
69	knee	0,21975	2,684	60,093	46,000
70	knuckle	0,00137	0,039	60,132	0,667
71	larynx	0,00410	0,078	60,210	1,333
72	leg	0,57612	4,979	65,189	85,333
73	limb	0,00500	0,039	65,228	0,667
74	lip	0,18584	2,100	67,328	36,000
75	liver	0,03991	0,700	68,028	12,000
76	lung	0,03181	0,545	68,573	9,333
77	mind	0,00912	0,078	68,650	1,333
78	mouth	0,50542	4,473	73,123	76,667
79	muscle	0,02318	0,467	73,590	8,000
80	nail	0,23941	2,800	76,391	48,000
81	navel	0,00077	0,039	76,429	0,667
82	neck	0,24475	2,606	79,035	44,667
83	nipple	0,00072	0,039	79,074	0,667
84	nose	0,53836	4,823	83,897	82,667
85	nostril	0,00880	0,078	83,975	1,333
86	palm	0,00537	0,039	84,014	0,667
87	pelvis	0,00433	0,039	84,053	0,667
88	penis	0,02487	0,350	84,403	6,000

89	physical	0,00667	0,039	84,442	0,667
90	pupil	0,00325	0,039	84,481	0,667
91	pussy	0,00631	0,078	84,559	1,333
92	shoulder	0,25409	2,489	87,048	42,667
93	skin	0,06846	0,817	87,865	14,000
94	skull	0,00990	0,117	87,981	2,000
95	spleen	0,01137	0,117	88,098	2,000
96	stomach	0,08103	1,245	89,343	21,333
97	stone	0,00248	0,078	89,420	1,333
98	teeth	0,24580	2,723	92,143	46,667
99	testicle	0,00137	0,039	92,182	0,667
100	thigh	0,00137	0,039	92,221	0,667
101	throat	0,01207	0,194	92,415	3,333
102	thumb	0,04352	0,506	92,921	8,667
103	tip	0,00302	0,039	92,960	0,667
104	tit	0,00111	0,039	92,999	0,667
105	toe	0,22448	2,606	95,605	44,667
106	tongue	0,15497	1,945	97,550	33,333
107	tooth	0,05488	0,622	98,172	10,667
108	triceps	0,00281	0,039	98,211	0,667
109	trunk	0,00500	0,039	98,250	0,667
110	tummy	0,00593	0,078	98,327	1,333
111	vagina	0,01262	0,233	98,561	4,000
112	vein	0,01147	0,233	98,794	4,000
113	waist	0,01077	0,194	98,989	3,333
114	wrinkle	0,00804	0,078	99,067	1,333
115	wrist	0,07670	0,933	100,000	16,000

2.1.2. Clothes

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	accessories	0,00465	0,048	0,048	0,667
2	anorak	0,04851	0,626	0,674	8,667
3	apron	0,01032	0,144	0,818	2,000
4	athletic shoe	0,00471	0,096	0,914	1,333
5	bag	0,02748	0,289	1,203	4,000
6	baseball cap	0,01164	0,144	1,347	2,000
7	bathing suit	0,00375	0,048	1,396	0,667
8	bathrobe	0,00414	0,096	1,492	1,333
9	bell-bottoms	0,00349	0,048	1,540	0,667
10	belt	0,08934	1,347	2,887	18,667
11	bikini	0,01221	0,192	3,080	2,667
12	blanket	0,00196	0,048	3,128	0,667
13	blazer	0,01103	0,144	3,272	2,000
14	blouse	0,11204	1,299	4,572	18,000
15	body	0,00183	0,048	4,620	0,667
16	body stocking	0,00465	0,048	4,668	0,667
17	bomber jacket	0,00325	0,048	4,716	0,667
18	Bonnet	0,00170	0,048	4,764	0,667
19	boot	0,17799	2,454	7,218	34,000
20	bow tie	0,00325	0,048	7,267	0,667
21	boxer shorts	0,03966	0,433	7,700	6,000
22	bra	0,15441	2,214	9,913	30,667
23	bracelet	0,00540	0,096	10,010	1,333
24	briefs	0,00158	0,048	10,058	0,667
25	button	0,01866	0,241	10,298	3,333
26	cap	0,06774	0,962	11,261	13,333
27	cardigan	0,08047	0,914	12,175	12,667
28	casual wear	0,00349	0,048	12,223	0,667
29	chemise	0,00537	0,048	12,271	0,667
30	chinos	0,00758	0,096	12,368	1,333
31	cloak	0,00281	0,048	12,416	0,667
32	clog	0,00096	0,048	12,464	0,667
33	coat	0,19665	2,454	14,918	34,000
34	cocktail dress	0,00211	0,048	14,966	0,667
35	collar	0,00227	0,048	15,014	0,667
36	cotton	0,01143	0,144	15,159	2,000
37	cuff	0,00103	0,048	15,207	0,667
38	cupboard	0,03002	0,385	15,592	5,333
39	designer	0,00158	0,048	15,640	0,667
40	dress	0,27882	3,272	18,912	45,333
41	dressing gown	0,00244	0,048	18,961	0,667
42	dungarees	0,00158	0,048	19,009	0,667
43	earring	0,00640	0,144	19,153	2,000
44	easy wear	0,00375	0,048	19,201	0,667
45	flat sandal	0,00183	0,048	19,249	0,667

46	fleece jacket	0,00111	0,048	19,297	0,667
47	flip-flop	0,01903	0,337	19,634	4,667
48	foot	0,00375	0,048	19,682	0,667
49	foulard	0,02118	0,289	19,971	4,000
50	French sneaker	0,00302	0,048	20,019	0,667
51	gaberdine	0,00211	0,048	20,067	0,667
52	glasses	0,02994	0,337	20,404	4,667
53	glove	0,17624	2,166	22,570	30,000
54	gown	0,00211	0,048	22,618	0,667
55	handbag	0,00375	0,048	22,666	0,667
56	handcuffs	0,00183	0,048	22,714	0,667
57	handkerchief	0,00586	0,096	22,810	1,333
58	hat	0,30655	3,465	26,275	48,000
59	headscarf	0,00325	0,048	26,323	0,667
60	heel	0,01859	0,385	26,708	5,333
61	high heels	0,02852	0,433	27,141	6,000
62	high-heeled shoe	0,00802	0,096	27,238	1,333
63	Hood	0,00665	0,096	27,334	1,333
64	hose	0,00620	0,048	27,382	0,667
65	hotpants	0,00281	0,048	27,430	0,667
66	jacket	0,34890	3,802	31,232	52,667
67	jeans	0,50265	4,909	36,141	68,000
68	jersey	0,09469	1,155	37,295	16,000
69	jumper	0,13067	1,203	38,499	16,667
70	jumpsuit	0,00432	0,096	38,595	1,333
71	knickers	0,04811	0,626	39,220	8,667
72	knit	0,00537	0,048	39,269	0,667
73	lace	0,00375	0,048	39,317	0,667
74	leggings	0,05682	0,770	40,087	10,667
75	lingerie	0,00647	0,096	40,183	1,333
76	mackintosh	0,00227	0,048	40,231	0,667
77	mini-short	0,01293	0,192	40,423	2,667
78	miniskirt	0,01613	0,241	40,664	3,333
79	mitten	0,00322	0,096	40,760	1,333
80	money	0,00262	0,048	40,808	0,667
81	necklace	0,01430	0,289	41,097	4,000
82	necktie	0,00500	0,048	41,145	0,667
83	nightdress	0,00281	0,048	41,193	0,667
84	nightgown	0,00605	0,096	41,290	1,333
85	overall	0,00740	0,144	41,434	2,000
86	panties	0,08272	1,107	42,541	15,333
87	pants	0,13574	1,492	44,033	20,667
88	parka	0,00137	0,048	44,081	0,667
89	pencil skirt	0,00375	0,048	44,129	0,667
90	pullover	0,19566	2,166	46,295	30,000
91	purse	0,00349	0,048	46,343	0,667
92	pyjamas	0,04720	0,770	47,113	10,667
93	raincoat	0,04274	0,722	47,834	10,000
94	ring	0,00807	0,144	47,979	2,000

95	robe	0,00752	0,096	48,075	1,333
96	rompers	0,00170	0,048	48,123	0,667
97	sandal	0,02175	0,289	48,412	4,000
98	scarf	0,23447	2,743	51,155	38,000
99	shirt	0,46558	4,379	55,534	60,667
100	shoe	0,53759	5,775	61,309	80,000
101	shop	0,00817	0,096	61,405	1,333
102	shopping centre	0,00577	0,048	61,453	0,667
103	short	0,00375	0,048	61,501	0,667
104	shorts	0,17067	2,021	63,523	28,000
105	shoulder pad	0,00537	0,048	63,571	0,667
106	skirt	0,55935	5,245	68,816	72,667
107	sleeper	0,00262	0,048	68,864	0,667
108	sleeve	0,00496	0,096	68,961	1,333
109	slip	0,05222	0,722	69,682	10,000
110	slipper	0,01832	0,337	70,019	4,667
111	Smart	0,00465	0,048	70,067	0,667
112	smoking jacket	0,01666	0,241	70,308	3,333
113	sneaker	0,01330	0,241	70,549	3,333
114	sock	0,40157	4,764	75,313	66,000
115	sports jacket	0,00806	0,096	75,409	1,333
116	stiletto	0,01418	0,192	75,602	2,667
117	stocking	0,01404	0,144	75,746	2,000
118	style	0,00147	0,048	75,794	0,667
119	suit	0,12693	1,684	77,478	23,333
120	sunglasses	0,01325	0,144	77,623	2,000
121	sweater	0,14538	1,636	79,259	22,667
122	sweatshirt	0,01042	0,096	79,355	1,333
123	swimming cap	0,00158	0,048	79,403	0,667
124	swimming costume	0,00409	0,096	79,500	1,333
125	swimming trunks	0,00465	0,048	79,548	0,667
126	swimsuit	0,00756	0,144	79,692	2,000
127	tennis shoe	0,00862	0,144	79,836	2,000
128	thong	0,01431	0,241	80,077	3,333
129	tie	0,13988	1,925	82,002	26,667
130	tights	0,02648	0,433	82,435	6,000
131	top	0,03200	0,433	82,868	6,000
132	towel	0,00137	0,048	82,916	0,667
133	tracksuit	0,04950	0,722	83,638	10,000
134	trainer	0,07278	0,914	84,552	12,667
135	trouser	0,71442	6,352	90,905	88,000
136	trunks	0,01023	0,144	91,049	2,000
137	T-shirt	0,66029	5,967	97,016	82,667
138	turtleneck	0,00564	0,096	97,113	1,333
139	tuxedo	0,00364	0,096	97,209	1,333
140	umbrella	0,01655	0,241	97,449	3,333
141	undergarment	0,00147	0,048	97,498	0,667
142	underpants	0,00808	0,096	97,594	1,333
143	underwear	0,13586	1,781	99,374	24,667

144	vest	0,01746	0,241	99,615	3,333
145	waist	0,00281	0,048	99,663	0,667
146	waistcoat	0,00281	0,048	99,711	0,667
147	watch	0,00700	0,096	99,808	1,333
148	wig	0,00403	0,048	99,856	0,667
149	wool	0,00500	0,048	99,904	0,667
150	wool shirt	0,00196	0,048	99,952	0,667
151	zip	0,00500	0,048	100,000	0,667

2.1.3. Food and drink

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	alcohol	0,03136	0,241	0,241	4,667
2	alcoholic drink	0,00553	0,034	0,275	0,667
3	alcopop	0,00788	0,103	0,378	2,000
4	almond	0,00297	0,034	0,413	0,667
5	anchovy	0,00431	0,034	0,447	0,667
6	Apple	0,20598	1,788	2,234	34,667
7	artichoke	0,00124	0,034	2,269	0,667
8	aubergine	0,01738	0,206	2,475	4,000
9	avocado	0,01334	0,172	2,647	3,333
10	bacon	0,01739	0,206	2,853	4,000
11	bagel	0,00160	0,034	2,888	0,667
12	baked potato	0,00381	0,034	2,922	0,667
13	banana	0,18259	1,650	4,572	32,000
14	bar	0,00262	0,034	4,606	0,667
15	barley	0,00358	0,034	4,641	0,667
16	BBQ	0,00405	0,034	4,675	0,667
17	bean	0,09278	0,859	5,535	16,667
18	beef	0,08628	0,791	6,325	15,333
19	beefsteak	0,01675	0,103	6,428	2,000
20	beer	0,24799	1,891	8,319	36,667
21	Berry	0,00694	0,069	8,388	1,333
22	biscuit	0,03009	0,275	8,663	5,333
23	black tea	0,00247	0,034	8,697	0,667
24	blackberry	0,00205	0,034	8,732	0,667
25	blackcurrant	0,00124	0,034	8,766	0,667
26	blueberry	0,00475	0,069	8,835	1,333
27	bottle	0,00081	0,034	8,869	0,667
28	brandy	0,00247	0,034	8,903	0,667
29	bread	0,24188	1,994	10,897	38,667
30	bread roll	0,00160	0,034	10,932	0,667
31	brisket	0,00110	0,034	10,966	0,667
32	broccoli	0,01101	0,138	11,103	2,667
33	Burger	0,03119	0,275	11,378	5,333
34	butter	0,07932	0,756	12,135	14,667
35	cabbage	0,01115	0,069	12,204	1,333
36	cake	0,11701	1,100	13,304	21,333
37	Candy	0,00582	0,103	13,407	2,000
38	cappuccino	0,00192	0,034	13,441	0,667
39	caramel	0,00160	0,034	13,475	0,667
40	carrot	0,13388	1,238	14,713	24,000
41	cauliflower	0,01541	0,172	14,885	3,333
42	cereal	0,01746	0,206	15,091	4,000
43	chamomile tea	0,00358	0,034	15,125	0,667
44	champagne	0,00431	0,034	15,160	0,667
45	cheese	0,09440	0,963	16,122	18,667

46	cheesecake	0,00279	0,034	16,157	0,667
47	cherry	0,01548	0,172	16,329	3,333
48	chestnut	0,00247	0,034	16,363	0,667
49	chicken	0,16837	1,581	17,944	30,667
50	chicken nugget	0,02146	0,138	18,082	2,667
51	chili	0,00667	0,034	18,116	0,667
52	Chinese food	0,00356	0,069	18,185	1,333
53	chip	0,22958	1,891	20,076	36,667
54	chocolate	0,09008	0,928	21,004	18,000
55	chop	0,01152	0,103	21,107	2,000
56	cider	0,01477	0,138	21,244	2,667
57	cinnamon	0,00402	0,069	21,313	1,333
58	clam	0,00232	0,034	21,348	0,667
59	Coca-Cola	0,05450	0,309	21,657	6,000
60	cocktail	0,00725	0,069	21,726	1,333
61	coconut	0,00799	0,069	21,794	1,333
62	cod	0,03105	0,309	22,104	6,000
63	codfish	0,00515	0,069	22,173	1,333
64	coffee	0,22335	2,097	24,270	40,667
65	cognac	0,00181	0,034	24,304	0,667
66	Coke	0,29469	2,200	26,504	42,667
67	cola	0,00553	0,034	26,538	0,667
68	coleslaw	0,00431	0,034	26,573	0,667
69	comfort food	0,00150	0,034	26,607	0,667
70	cookie	0,02821	0,344	26,951	6,667
71	cordial	0,00297	0,034	26,985	0,667
72	corn	0,01196	0,172	27,157	3,333
73	cornflakes	0,01144	0,103	27,260	2,000
74	cottage pie	0,00205	0,034	27,295	0,667
75	cow	0,00405	0,034	27,329	0,667
76	crab	0,00459	0,034	27,363	0,667
77	cranberry	0,00589	0,103	27,466	2,000
78	cream	0,00426	0,069	27,535	1,333
79	crème caramel	0,00170	0,034	27,570	0,667
80	crepe	0,00141	0,034	27,604	0,667
81	crisp	0,00763	0,069	27,673	1,333
82	cucumber	0,03657	0,378	28,051	7,333
83	cup	0,00297	0,034	28,085	0,667
84	cupcake	0,00925	0,069	28,154	1,333
85	dessert	0,01312	0,138	28,292	2,667
86	Diet Coke	0,00613	0,069	28,360	1,333
87	drink	0,01072	0,069	28,429	1,333
88	duck	0,00835	0,069	28,498	1,333
89	egg	0,14981	1,169	29,667	22,667
90	eggplant	0,00899	0,103	29,770	2,000
91	energy drink	0,00358	0,034	29,804	0,667
92	Fanta	0,00997	0,103	29,907	2,000
93	fast food	0,01007	0,069	29,976	1,333
94	fig	0,00297	0,034	30,010	0,667

95	fish	0,39906	3,128	33,139	60,667
96	fish and chips	0,00891	0,103	33,242	2,000
97	flour	0,00816	0,069	33,310	1,333
98	foie gras	0,00520	0,034	33,345	0,667
99	fork	0,00336	0,034	33,379	0,667
100	French fries	0,00628	0,069	33,448	1,333
101	fresh milk	0,00431	0,034	33,482	0,667
102	fried	0,00738	0,069	33,551	1,333
103	fruit	0,11622	0,928	34,479	18,000
104	fruit juice	0,00489	0,034	34,514	0,667
105	garlic	0,03295	0,413	34,926	8,000
106	gazpacho	0,00520	0,034	34,960	0,667
107	gin	0,02005	0,172	35,132	3,333
108	gin and tonic	0,00667	0,034	35,167	0,667
109	ginger	0,00489	0,034	35,201	0,667
110	ginger ale	0,00232	0,034	35,235	0,667
111	glass	0,00547	0,103	35,339	2,000
112	gooseberry	0,00218	0,034	35,373	0,667
113	grab	0,00381	0,034	35,407	0,667
114	grape	0,03192	0,344	35,751	6,667
115	grapefruit	0,00431	0,034	35,785	0,667
116	grapefruit juice	0,00358	0,034	35,820	0,667
117	gravy	0,00381	0,034	35,854	0,667
118	green tea	0,00205	0,034	35,889	0,667
119	hake	0,00459	0,034	35,923	0,667
120	ham	0,03187	0,378	36,301	7,333
121	hamburger	0,18764	1,409	37,711	27,333
122	honey	0,01190	0,103	37,814	2,000
123	hot chocolate	0,00405	0,034	37,848	0,667
124	hot dog	0,01137	0,138	37,986	2,667
125	hum	0,00124	0,034	38,020	0,667
126	ice	0,00873	0,069	38,089	1,333
127	ice cream	0,04434	0,550	38,639	10,667
128	Indian food	0,00667	0,034	38,673	0,667
129	jam	0,05529	0,619	39,292	12,000
130	jelly	0,00489	0,034	39,326	0,667
131	joint	0,00520	0,034	39,361	0,667
132	juice	0,25189	2,166	41,526	42,000
133	ketchup	0,01518	0,138	41,664	2,667
134	Kir	0,00218	0,034	41,698	0,667
135	Kir Royal	0,00205	0,034	41,733	0,667
136	kiwi	0,00969	0,103	41,836	2,000
137	lamb	0,04606	0,481	42,317	9,333
138	leek	0,00132	0,034	42,351	0,667
139	legume	0,00336	0,034	42,386	0,667
140	lemon	0,13140	1,203	43,589	23,333
141	lemonade	0,04331	0,447	44,036	8,667
142	lentil	0,01168	0,103	44,139	2,000
143	lettuce	0,14051	1,375	45,514	26,667

144	light	0,00316	0,034	45,548	0,667
145	liquor	0,02208	0,206	45,755	4,000
146	main course	0,00141	0,034	45,789	0,667
147	mandarin	0,00337	0,069	45,858	1,333
148	marmalade	0,01387	0,138	45,995	2,667
149	mayonnaise	0,01101	0,138	46,133	2,667
150	McDonald's	0,00667	0,034	46,167	0,667
151	meal	0,02917	0,172	46,339	3,333
152	meat	0,40602	2,888	49,227	56,000
153	meat loaf	0,00117	0,034	49,261	0,667
154	meatball	0,02811	0,241	49,502	4,667
155	melón	0,04754	0,584	50,086	11,333
156	milk	0,29564	2,372	52,458	46,000
157	milkshake	0,02776	0,275	52,733	5,333
158	mineral water	0,00553	0,034	52,767	0,667
159	mojito	0,00806	0,069	52,836	1,333
160	monkfish	0,00626	0,034	52,870	0,667
161	muffin	0,00205	0,034	52,905	0,667
162	mushroom	0,01082	0,138	53,042	2,667
163	mussel	0,00358	0,034	53,077	0,667
164	mustard	0,00086	0,034	53,111	0,667
165	nacho	0,00489	0,034	53,145	0,667
166	noodle	0,00685	0,069	53,214	1,333
167	nut	0,00963	0,069	53,283	1,333
168	oat	0,00431	0,034	53,317	0,667
169	octopus	0,01008	0,138	53,455	2,667
170	oil	0,08142	0,963	54,417	18,667
171	olive	0,01739	0,206	54,624	4,000
172	olive oil	0,01034	0,138	54,761	2,667
173	omelette	0,04776	0,378	55,139	7,333
174	onion	0,13747	1,238	56,377	24,000
175	orange	0,28796	2,338	58,714	45,333
176	orange juice	0,06745	0,584	59,299	11,333
177	oxtail soup	0,00553	0,034	59,333	0,667
178	paella	0,02583	0,172	59,505	3,333
179	pancake	0,00232	0,034	59,539	0,667
180	panna cotta	0,00160	0,034	59,574	0,667
181	paprika	0,00192	0,034	59,608	0,667
182	pasta	0,08330	0,722	60,330	14,000
183	paste	0,00358	0,034	60,364	0,667
184	pea	0,03022	0,378	60,743	7,333
185	peach	0,02388	0,309	61,052	6,000
186	peanut	0,01001	0,138	61,189	2,667
187	pear	0,06695	0,688	61,877	13,333
188	penne	0,00626	0,034	61,911	0,667
189	pepper	0,09183	0,997	62,908	19,333
190	Pepsi	0,00873	0,069	62,977	1,333
191	pie	0,02888	0,309	63,286	6,000
192	pig	0,00694	0,069	63,355	1,333

193	pineapple	0,08196	0,756	64,111	14,667
194	pizza	0,16227	1,272	65,383	24,667
195	plate	0,00779	0,103	65,486	2,000
196	plum	0,00170	0,034	65,521	0,667
197	plum pudding	0,00316	0,034	65,555	0,667
198	popcorn	0,01202	0,138	65,693	2,667
199	pork	0,06601	0,688	66,380	13,333
200	pork chop	0,00959	0,069	66,449	1,333
201	potato	0,32813	2,750	69,199	53,333
202	potato crisp	0,01400	0,103	69,302	2,000
203	poultry	0,00297	0,034	69,337	0,667
204	prawn	0,01363	0,172	69,508	3,333
205	pudding	0,00489	0,034	69,543	0,667
206	pumpkin	0,00805	0,069	69,612	1,333
207	rabbit	0,00786	0,069	69,680	1,333
208	rare	0,00358	0,034	69,715	0,667
209	raspberry	0,00520	0,034	69,749	0,667
210	ravioli	0,00297	0,034	69,783	0,667
211	red tea	0,00232	0,034	69,818	0,667
212	red wine	0,01277	0,138	69,955	2,667
213	refresher	0,00626	0,034	69,990	0,667
214	restaurant	0,00757	0,103	70,093	2,000
215	rib	0,00316	0,034	70,127	0,667
216	rib eye	0,00336	0,034	70,162	0,667
217	rice	0,15020	1,169	71,330	22,667
218	roast	0,01097	0,103	71,433	2,000
219	roast beef	0,00531	0,069	71,502	1,333
220	roast chicken	0,00934	0,069	71,571	1,333
221	rum	0,03460	0,344	71,915	6,667
222	salad	0,14705	1,375	73,290	26,667
223	salami	0,00247	0,034	73,324	0,667
224	salmon	0,01340	0,172	73,496	3,333
225	salt	0,10292	1,169	74,665	22,667
226	sandwich	0,10083	0,859	75,524	16,667
227	sangría	0,00205	0,034	75,559	0,667
228	sardine	0,00405	0,034	75,593	0,667
229	sauce	0,04783	0,447	76,040	8,667
230	sausage	0,06795	0,722	76,762	14,000
231	scampi	0,00160	0,034	76,796	0,667
232	seafood	0,02215	0,206	77,002	4,000
233	shellfish	0,00553	0,034	77,037	0,667
234	shrimp	0,00405	0,034	77,071	0,667
235	smoothie	0,00247	0,034	77,106	0,667
236	snack	0,01788	0,138	77,243	2,667
237	snail	0,00649	0,069	77,312	1,333
238	soda	0,05878	0,447	77,759	8,667
239	soft drink	0,02393	0,206	77,965	4,000
240	sole	0,00751	0,069	78,034	1,333
241	soup	0,13753	1,169	79,202	22,667

242	spaghetti	0,08310	0,584	79,787	11,333
243	sparkling water	0,00562	0,069	79,856	1,333
244	sparkling wine	0,00431	0,034	79,890	0,667
245	spice	0,00358	0,034	79,924	0,667
246	spirit	0,02774	0,241	80,165	4,667
247	spoon	0,00358	0,034	80,199	0,667
248	spring water	0,00279	0,034	80,234	0,667
249	squid	0,00714	0,103	80,337	2,000
250	starter	0,00150	0,034	80,371	0,667
251	steak	0,07124	0,619	80,990	12,000
252	stew	0,01710	0,172	81,162	3,333
253	still water	0,00181	0,034	81,196	0,667
254	strawberry	0,09864	1,031	82,228	20,000
255	sugar	0,10577	1,134	83,362	22,000
256	sunflower seed oil	0,00170	0,034	83,396	0,667
257	sushi	0,00653	0,069	83,465	1,333
258	sweet	0,01985	0,275	83,740	5,333
259	table	0,00218	0,034	83,774	0,667
260	take away	0,01077	0,069	83,843	1,333
261	tangerine	0,00589	0,034	83,878	0,667
262	tea	0,15251	1,444	85,321	28,000
263	toast	0,01209	0,138	85,459	2,667
264	tomato	0,35255	3,163	88,622	61,333
265	tomato sauce	0,00218	0,034	88,656	0,667
266	tortellini	0,00316	0,034	88,690	0,667
267	traditional food	0,00553	0,034	88,725	0,667
268	tuna	0,05059	0,619	89,343	12,000
269	turkey	0,00336	0,034	89,378	0,667
270	turnip	0,00489	0,034	89,412	0,667
271	veal	0,00613	0,069	89,481	1,333
272	vegetable	0,19399	1,409	90,890	27,333
273	veggie	0,00205	0,034	90,925	0,667
274	vinegar	0,02119	0,275	91,200	5,333
275	vodka	0,01633	0,172	91,372	3,333
276	walnut	0,00262	0,034	91,406	0,667
277	wasabi	0,00297	0,034	91,440	0,667
278	wáter	0,56424	3,884	95,325	75,333
279	watermelon	0,08564	0,859	96,184	16,667
280	weir	0,00091	0,034	96,219	0,667
281	well-cooked	0,00381	0,034	96,253	0,667
282	whisky	0,08990	0,825	97,078	16,000
283	white tea	0,00218	0,034	97,112	0,667
284	white wine	0,01218	0,103	97,216	2,000
285	wine	0,30385	2,303	99,519	44,667
286	yogurt	0,03903	0,447	99,966	8,667
287	Yorkshire pudding	0,00405	0,034	100,000	0,667

2.1.4. The school (furniture and school material)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	adolescent	0,00577	0,047	0,047	0,667
2	agenda	0,00440	0,094	0,141	1,333
3	armchair	0,00782	0,094	0,235	1,333
4	art college	0,00183	0,047	0,282	0,667
5	assessment	0,00170	0,047	0,329	0,667
6	backpack	0,01216	0,141	0,471	2,000
7	bag	0,09744	1,271	1,741	18,000
8	ball	0,00606	0,094	1,835	1,333
9	ballpoint pen	0,05836	0,612	2,447	8,667
10	bed	0,00302	0,047	2,494	0,667
11	bell	0,00403	0,047	2,541	0,667
12	bench	0,00819	0,094	2,635	1,333
13	bin	0,03645	0,471	3,106	6,667
14	biology	0,00325	0,047	3,153	0,667
15	blackboard	0,47853	4,376	7,529	62,000
16	blind	0,00781	0,141	7,671	2,000
17	blog	0,00349	0,047	7,718	0,667
18	board	0,06466	0,706	8,424	10,000
19	book	0,48006	5,271	13,694	74,667
20	bookcase	0,00302	0,047	13,741	0,667
21	bookshelf	0,00375	0,047	13,788	0,667
22	boot	0,00500	0,047	13,835	0,667
23	box	0,00375	0,047	13,882	0,667
24	break	0,01885	0,282	14,165	4,000
25	cable	0,00281	0,047	14,212	0,667
26	cafetería	0,00227	0,047	14,259	0,667
27	calculator	0,01810	0,282	14,541	4,000
28	carpet	0,02581	0,329	14,871	4,667
29	case	0,00196	0,047	14,918	0,667
30	cassette	0,00349	0,047	14,965	0,667
31	CD player	0,01139	0,141	15,106	2,000
32	ceiling	0,00513	0,094	15,200	1,333
33	chair	0,64510	5,788	20,988	82,000
34	chalk	0,14936	1,835	22,824	26,000
35	cheat	0,00577	0,047	22,871	0,667
36	child	0,02018	0,188	23,059	2,667
37	children	0,01364	0,188	23,247	2,667
38	class	0,03795	0,424	23,671	6,000
39	class list	0,01043	0,094	23,765	1,333
40	classmate	0,00821	0,094	23,859	1,333
41	classroom	0,09871	1,129	24,988	16,000
42	clip	0,00325	0,047	25,035	0,667
43	clock	0,00196	0,047	25,082	0,667
44	closet	0,00927	0,094	25,176	1,333
45	color	0,03395	0,424	25,600	6,000

46	compass	0,00262	0,047	25,647	0,667
47	computer	0,30757	3,953	29,600	56,000
48	concierge	0,00262	0,047	29,647	0,667
49	copy	0,00674	0,094	29,741	1,333
50	cork	0,00262	0,047	29,788	0,667
51	correcting fluid	0,00302	0,047	29,835	0,667
52	corridor	0,00196	0,047	29,882	0,667
53	coulored pencil	0,00183	0,047	29,929	0,667
54	course	0,00147	0,047	29,976	0,667
55	coursebook	0,00433	0,047	30,024	0,667
56	crayón	0,01284	0,188	30,212	2,667
57	cupboard	0,01871	0,235	30,447	3,333
58	cutter	0,00302	0,047	30,494	0,667
59	deadline	0,00158	0,047	30,541	0,667
60	dean	0,00465	0,047	30,588	0,667
61	desk	0,31190	2,824	33,412	40,000
62	desktop computer	0,00500	0,047	33,459	0,667
63	diary	0,00183	0,047	33,506	0,667
64	dictionary	0,05989	0,847	34,353	12,000
65	digital board	0,00727	0,094	34,447	1,333
66	dining hall	0,00244	0,047	34,494	0,667
67	door	0,13374	1,694	36,188	24,000
68	draw	0,00757	0,141	36,329	2,000
69	drawing	0,00465	0,047	36,376	0,667
70	dull	0,00537	0,047	36,424	0,667
71	dust cover	0,00375	0,047	36,471	0,667
72	DVD	0,00354	0,094	36,565	1,333
73	DVD player	0,00302	0,047	36,612	0,667
74	encyclopedia	0,00349	0,047	36,659	0,667
75	English	0,00375	0,047	36,706	0,667
76	English lesson	0,00899	0,094	36,800	1,333
77	entrance hall	0,00302	0,047	36,847	0,667
78	eraser	0,09972	1,318	38,165	18,667
79	essay	0,00183	0,047	38,212	0,667
80	exam	0,04219	0,565	38,776	8,000
81	exercise	0,01246	0,188	38,965	2,667
82	faculty	0,00433	0,047	39,012	0,667
83	fail	0,00158	0,047	39,059	0,667
84	file	0,00873	0,141	39,200	2,000
85	flashcard	0,00375	0,047	39,247	0,667
86	floor	0,00606	0,094	39,341	1,333
87	flute	0,00196	0,047	39,388	0,667
88	folder	0,04611	0,612	40,000	8,667
89	football match	0,00349	0,047	40,047	0,667
90	friend	0,00665	0,094	40,141	1,333
91	furniture	0,00667	0,047	40,188	0,667
92	game	0,00444	0,094	40,282	1,333
93	garden	0,00465	0,047	40,329	0,667
94	glasses	0,00644	0,094	40,424	1,333

95	glue	0,01044	0,141	40,565	2,000
96	grammar	0,00537	0,047	40,612	0,667
97	gym	0,05160	0,706	41,318	10,000
98	gymnastics	0,01262	0,141	41,459	2,000
99	hall	0,00933	0,094	41,553	1,333
100	hanger	0,02516	0,329	41,882	4,667
101	headmaster	0,00715	0,094	41,976	1,333
102	heat	0,00487	0,094	42,071	1,333
103	hi-fi	0,00403	0,047	42,118	0,667
104	high school	0,01103	0,141	42,259	2,000
105	highlighter	0,00888	0,188	42,447	2,667
106	history	0,00752	0,094	42,541	1,333
107	homework	0,01331	0,188	42,729	2,667
108	Internet	0,01356	0,141	42,871	2,000
109	Internet access	0,00302	0,047	42,918	0,667
110	keyboard	0,00683	0,094	43,012	1,333
111	kindergarten	0,00684	0,094	43,106	1,333
112	knowledge	0,00262	0,047	43,153	0,667
113	laboratory	0,01457	0,282	43,435	4,000
114	lamp	0,01880	0,329	43,765	4,667
115	language	0,00089	0,047	43,812	0,667
116	laptop	0,01266	0,188	44,000	2,667
117	leaf	0,00433	0,047	44,047	0,667
118	learn	0,00600	0,094	44,141	1,333
119	lesson	0,01229	0,141	44,282	2,000
120	library	0,02404	0,329	44,612	4,667
121	lift	0,00577	0,047	44,659	0,667
122	light	0,01709	0,188	44,847	2,667
123	light bulb	0,00465	0,047	44,894	0,667
124	litter	0,00196	0,047	44,941	0,667
125	map	0,07756	1,082	46,024	15,333
126	mark	0,01309	0,188	46,212	2,667
127	marker	0,02105	0,329	46,541	4,667
128	master	0,00465	0,047	46,588	0,667
129	master class	0,00325	0,047	46,635	0,667
130	mathematics	0,00903	0,094	46,729	1,333
131	maths	0,02682	0,282	47,012	4,000
132	microphone	0,00349	0,047	47,059	0,667
133	mirror	0,00620	0,047	47,106	0,667
134	mouse	0,01868	0,329	47,435	4,667
135	music	0,00537	0,047	47,482	0,667
136	netbook	0,01424	0,188	47,671	2,667
137	network	0,00375	0,047	47,718	0,667
138	newspaper	0,00244	0,047	47,765	0,667
139	note	0,01873	0,235	48,000	3,333
140	notebook	0,30747	3,624	51,624	51,333
141	notepad	0,02028	0,282	51,906	4,000
142	notice board	0,00537	0,047	51,953	0,667
143	office	0,00281	0,047	52,000	0,667

144	oral examination	0,00147	0,047	52,047	0,667
145	overhead projector	0,00433	0,047	52,094	0,667
146	page	0,01896	0,188	52,282	2,667
147	paint	0,01235	0,188	52,471	2,667
148	panel	0,00375	0,047	52,518	0,667
149	paper	0,20505	2,447	54,965	34,667
150	paper sheet	0,01043	0,094	55,059	1,333
151	parking	0,00158	0,047	55,106	0,667
152	partner	0,00375	0,047	55,153	0,667
153	pass	0,00170	0,047	55,200	0,667
154	pavement	0,00500	0,047	55,247	0,667
155	PC	0,01009	0,141	55,388	2,000
156	pen	0,54858	5,224	60,612	74,000
157	pencil	0,75392	6,965	67,576	98,667
158	pencil case	0,08178	1,082	68,659	15,333
159	pencil sharpener	0,00631	0,094	68,753	1,333
160	people	0,00375	0,047	68,800	0,667
161	personal computer	0,00302	0,047	68,847	0,667
162	philosophy	0,00375	0,047	68,894	0,667
163	physics	0,00405	0,094	68,988	1,333
164	picture	0,00685	0,094	69,082	1,333
165	piece of paper	0,00500	0,047	69,129	0,667
166	pin	0,00281	0,047	69,176	0,667
167	plant	0,00170	0,047	69,224	0,667
168	Plasticine	0,00196	0,047	69,271	0,667
169	Play-Doh	0,00183	0,047	69,318	0,667
170	playground	0,02688	0,329	69,647	4,667
171	plug	0,00262	0,047	69,694	0,667
172	pool	0,00211	0,047	69,741	0,667
173	portfolio	0,00465	0,047	69,788	0,667
174	poster	0,01869	0,235	70,024	3,333
175	Post-it	0,01034	0,141	70,165	2,000
176	primary school	0,00302	0,047	70,212	0,667
177	principal	0,00508	0,094	70,306	1,333
178	professor	0,02844	0,282	70,588	4,000
179	projector	0,04271	0,518	71,106	7,333
180	puncher	0,00158	0,047	71,153	0,667
181	punishment	0,00302	0,047	71,200	0,667
182	pupil	0,10986	1,224	72,424	17,333
183	racket	0,00302	0,047	72,471	0,667
184	radio	0,01275	0,188	72,659	2,667
185	read	0,00227	0,047	72,706	0,667
186	roof	0,00325	0,047	72,753	0,667
187	rubber	0,29226	3,294	76,047	46,667
188	rubbish	0,00170	0,047	76,094	0,667
189	rucksack	0,00855	0,141	76,235	2,000
190	rule	0,12303	1,365	77,600	19,333
191	ruler	0,14245	1,553	79,153	22,000
192	sandwich	0,00403	0,047	79,200	0,667

193	school	0,00667	0,047	79,247	0,667
194	schoolbag	0,03998	0,612	79,859	8,667
195	science	0,00325	0,047	79,906	0,667
196	scissors	0,01121	0,188	80,094	2,667
197	screen	0,02532	0,329	80,424	4,667
198	screwdriver	0,00403	0,047	80,471	0,667
199	seat	0,01198	0,094	80,565	1,333
200	Sellotape	0,00170	0,047	80,612	0,667
201	sharpener	0,03658	0,471	81,082	6,667
202	sheet	0,05139	0,612	81,694	8,667
203	sheet of paper	0,00302	0,047	81,741	0,667
204	shelf	0,01759	0,329	82,071	4,667
205	shirt	0,00465	0,047	82,118	0,667
206	shoulder bag	0,00325	0,047	82,165	0,667
207	sofá	0,00349	0,047	82,212	0,667
208	speaker	0,01063	0,141	82,353	2,000
209	sponge	0,00302	0,047	82,400	0,667
210	stair	0,01264	0,141	82,541	2,000
211	staple	0,00128	0,047	82,588	0,667
212	stapler	0,01075	0,188	82,776	2,667
213	stereo	0,00325	0,047	82,824	0,667
214	sticker	0,00211	0,047	82,871	0,667
215	student	0,10003	1,176	84,047	16,667
216	student book	0,02249	0,329	84,376	4,667
217	study	0,01006	0,141	84,518	2,000
218	subject	0,02924	0,424	84,941	6,000
219	suitcase	0,01122	0,141	85,082	2,000
220	table	0,66936	5,741	90,824	81,333
221	task	0,00665	0,094	90,918	1,333
222	teacher	0,25663	2,682	93,600	38,000
223	teenager	0,01003	0,094	93,694	1,333
224	television	0,03150	0,424	94,118	6,000
225	tennis	0,00433	0,047	94,165	0,667
226	test	0,01409	0,188	94,353	2,667
227	text	0,00465	0,047	94,400	0,667
228	textbook	0,01812	0,188	94,588	2,667
229	timetable	0,03071	0,376	94,965	5,333
230	toilet	0,02184	0,282	95,247	4,000
231	uncomfortable	0,00577	0,047	95,294	0,667
232	university	0,00778	0,094	95,388	1,333
233	vending machine	0,00537	0,047	95,435	0,667
234	video	0,00678	0,094	95,529	1,333
235	Wall	0,01391	0,141	95,671	2,000
236	wardrobe	0,05209	0,659	96,329	9,333
237	waste paper	0,00147	0,047	96,376	0,667
238	WC	0,00500	0,047	96,424	0,667
239	whiteboard	0,02916	0,282	96,706	4,000
240	window	0,18051	2,259	98,965	32,000
241	Wood	0,00433	0,047	99,012	0,667

242	workbook	0,07543	0,894	99,906	12,667
243	write	0,00244	0,047	99,953	0,667
244	written examination	0,00137	0,047	100,000	0,667

2.1.5. The city

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	accident	0,00537	0,043	0,043	0,667
2	advertising	0,00631	0,085	0,128	1,333
3	airplane	0,00465	0,043	0,171	0,667
4	airport	0,07504	0,938	1,108	14,667
5	alley	0,00403	0,043	1,151	0,667
6	ambulance	0,00637	0,085	1,236	1,333
7	animal	0,00375	0,043	1,279	0,667
8	apartment	0,01003	0,085	1,364	1,333
9	asphalt	0,00244	0,043	1,407	0,667
10	avenue	0,08465	0,895	2,302	14,000
11	baker's shop	0,00325	0,043	2,344	0,667
12	bank	0,09605	1,108	3,453	17,333
13	bar	0,08270	0,938	4,390	14,667
14	barber's shop	0,00158	0,043	4,433	0,667
15	beach	0,02869	0,298	4,731	4,667
16	Benalmádena	0,00620	0,043	4,774	0,667
17	bench	0,01495	0,213	4,987	3,333
18	bicycle	0,04658	0,512	5,499	8,000
19	bike	0,03900	0,384	5,882	6,000
20	billboard	0,00537	0,043	5,925	0,667
21	bin	0,00465	0,043	5,968	0,667
22	bingo	0,00403	0,043	6,010	0,667
23	bird	0,00262	0,043	6,053	0,667
24	block	0,00196	0,043	6,095	0,667
25	book store	0,00340	0,085	6,181	1,333
26	bookshop	0,02062	0,256	6,436	4,000
27	boulevard	0,01455	0,171	6,607	2,667
28	bridge	0,04454	0,512	7,118	8,000
29	building	0,43831	3,623	10,742	56,667
30	bump	0,00667	0,043	10,784	0,667
31	bus	0,29564	2,728	13,512	42,667
32	bus station	0,05945	0,639	14,152	10,000
33	bus stop	0,08059	1,066	15,217	16,667
34	butcher's shop	0,01383	0,171	15,388	2,667
35	butchery	0,00281	0,043	15,431	0,667
36	cab	0,00926	0,128	15,558	2,000
37	café	0,01274	0,171	15,729	2,667
38	cafetería	0,01463	0,171	15,899	2,667
39	campus	0,00170	0,043	15,942	0,667
40	capital city	0,00667	0,043	15,985	0,667
41	car	0,54884	4,561	20,546	71,333
42	car park	0,01282	0,171	20,716	2,667
43	carpenter	0,00147	0,043	20,759	0,667
44	casino	0,00433	0,043	20,801	0,667
45	castle	0,02607	0,213	21,014	3,333

46	cat	0,01517	0,171	21,185	2,667
47	catedral	0,05555	0,639	21,824	10,000
48	cemetery	0,00147	0,043	21,867	0,667
49	centre	0,03173	0,298	22,165	4,667
50	chemist's shop	0,01352	0,128	22,293	2,000
51	chemistry laboratory	0,00325	0,043	22,336	0,667
52	child	0,00500	0,043	22,379	0,667
53	children	0,01560	0,256	22,634	4,000
54	church	0,11037	1,066	23,700	16,667
55	cinema	0,12876	1,449	25,149	22,667
56	circus	0,01167	0,085	25,234	1,333
57	citadel	0,01209	0,128	25,362	2,000
58	citizen	0,01891	0,213	25,575	3,333
59	city	0,00667	0,043	25,618	0,667
60	city centre	0,06255	0,639	26,257	10,000
61	city hall	0,04904	0,512	26,769	8,000
62	clinic	0,00211	0,043	26,812	0,667
63	clothes shop	0,00137	0,043	26,854	0,667
64	club	0,00902	0,085	26,939	1,333
65	coffee	0,00403	0,043	26,982	0,667
66	coffee shop	0,00349	0,043	27,025	0,667
67	commercial centre	0,02755	0,341	27,366	5,333
68	company	0,00620	0,043	27,408	0,667
69	concert	0,00819	0,085	27,494	1,333
70	contamination	0,00778	0,085	27,579	1,333
71	corner	0,00281	0,043	27,621	0,667
72	corruption	0,00929	0,085	27,707	1,333
73	cottage	0,00244	0,043	27,749	0,667
74	council	0,01690	0,213	27,962	3,333
75	crew	0,00465	0,043	28,005	0,667
76	crime	0,00563	0,085	28,090	1,333
77	crowd	0,00545	0,085	28,176	1,333
78	culture	0,00302	0,043	28,218	0,667
79	cycle lane	0,00433	0,043	28,261	0,667
80	danger	0,00433	0,043	28,303	0,667
81	department store	0,00325	0,043	28,346	0,667
82	detached house	0,00211	0,043	28,389	0,667
83	dirt	0,00537	0,043	28,431	0,667
84	disco(theque)	0,02581	0,256	28,687	4,000
85	discussion	0,00375	0,043	28,730	0,667
86	diversión	0,00471	0,085	28,815	1,333
87	dog	0,03261	0,384	29,199	6,000
88	door	0,00433	0,043	29,241	0,667
89	double-decker bus	0,00281	0,043	29,284	0,667
90	downtown	0,01008	0,128	29,412	2,000
91	Dublin	0,00375	0,043	29,454	0,667
92	dustbin	0,00244	0,043	29,497	0,667
93	earth	0,00537	0,043	29,540	0,667
94	electricity	0,00537	0,043	29,582	0,667

95	England	0,00325	0,043	29,625	0,667
96	facade	0,00281	0,043	29,668	0,667
97	fashion	0,00196	0,043	29,710	0,667
98	fence	0,00325	0,043	29,753	0,667
99	fire station	0,00433	0,043	29,795	0,667
100	firefighter	0,00325	0,043	29,838	0,667
101	fireman	0,00649	0,128	29,966	2,000
102	fish shop	0,00403	0,043	30,009	0,667
103	fishmonger	0,00325	0,043	30,051	0,667
104	flat	0,11041	1,108	31,159	17,333
105	floor	0,00325	0,043	31,202	0,667
106	flower	0,00634	0,128	31,330	2,000
107	flower shop	0,00375	0,043	31,373	0,667
108	fortress	0,00695	0,085	31,458	1,333
109	fountain	0,03503	0,384	31,841	6,000
110	friend	0,00698	0,085	31,927	1,333
111	Fuengirola	0,00537	0,043	31,969	0,667
112	garaje	0,00262	0,043	32,012	0,667
113	garden	0,07590	0,895	32,907	14,000
114	government	0,00244	0,043	32,950	0,667
115	grass	0,00577	0,043	32,992	0,667
116	graveyard	0,00465	0,043	33,035	0,667
117	grocery store	0,00158	0,043	33,078	0,667
118	ground	0,00537	0,043	33,120	0,667
119	gym	0,00988	0,171	33,291	2,667
120	gypsy	0,00358	0,085	33,376	1,333
121	hairdresser	0,00540	0,085	33,461	1,333
122	hall	0,00379	0,085	33,546	1,333
123	harbour	0,01190	0,128	33,674	2,000
124	hedge	0,00349	0,043	33,717	0,667
125	helicopter	0,00403	0,043	33,760	0,667
126	high school	0,01581	0,171	33,930	2,667
127	high Street	0,00196	0,043	33,973	0,667
128	highway	0,02343	0,298	34,271	4,667
129	hill	0,00465	0,043	34,314	0,667
130	history	0,00111	0,043	34,356	0,667
131	home	0,00620	0,043	34,399	0,667
132	homelessness	0,00586	0,085	34,484	1,333
133	hospital	0,08028	0,853	35,337	13,333
134	hostel	0,00227	0,043	35,379	0,667
135	hotel	0,03070	0,426	35,806	6,667
136	house	0,24468	2,344	38,150	36,667
137	hurry	0,00147	0,043	38,193	0,667
138	hypermarket	0,00281	0,043	38,235	0,667
139	inhabitant	0,00349	0,043	38,278	0,667
140	institution	0,00349	0,043	38,321	0,667
141	intersection	0,00433	0,043	38,363	0,667
142	Irish pub	0,00375	0,043	38,406	0,667
143	job	0,00375	0,043	38,448	0,667

144	joy	0,00262	0,043	38,491	0,667
145	junction	0,00244	0,043	38,534	0,667
146	kiosk	0,00211	0,043	38,576	0,667
147	lake	0,00670	0,085	38,662	1,333
148	lamp	0,00698	0,085	38,747	1,333
149	lamppost	0,00196	0,043	38,789	0,667
150	landscape	0,00667	0,043	38,832	0,667
151	language	0,00325	0,043	38,875	0,667
152	laundry	0,00695	0,085	38,960	1,333
153	library	0,06166	0,767	39,727	12,000
154	lift	0,01155	0,085	39,812	1,333
155	light	0,06585	0,554	40,367	8,667
156	lighthouse	0,00349	0,043	40,409	0,667
157	litter	0,00584	0,085	40,494	1,333
158	litter bin	0,00489	0,085	40,580	1,333
159	London	0,01860	0,171	40,750	2,667
160	luxury	0,00302	0,043	40,793	0,667
161	Madrid	0,00147	0,043	40,835	0,667
162	mail	0,00375	0,043	40,878	0,667
163	mailbox	0,00525	0,085	40,963	1,333
164	main road	0,00375	0,043	41,006	0,667
165	main Street	0,00403	0,043	41,049	0,667
166	main town	0,00183	0,043	41,091	0,667
167	main way	0,00465	0,043	41,134	0,667
168	major	0,00667	0,043	41,176	0,667
169	Málaga	0,01924	0,171	41,347	2,667
170	mal	0,05079	0,767	42,114	12,000
171	man	0,00262	0,043	42,157	0,667
172	mansión	0,00196	0,043	42,199	0,667
173	map	0,01102	0,128	42,327	2,000
174	Marbella	0,00500	0,043	42,370	0,667
175	market	0,04407	0,597	42,967	9,333
176	market place	0,00211	0,043	43,009	0,667
177	mayor	0,00262	0,043	43,052	0,667
178	meadow	0,00128	0,043	43,095	0,667
179	metro	0,03173	0,298	43,393	4,667
180	metropolitan	0,01077	0,085	43,478	1,333
181	Mijas	0,00577	0,043	43,521	0,667
182	money	0,00477	0,085	43,606	1,333
183	monolith	0,00281	0,043	43,649	0,667
184	monument	0,08942	0,853	44,501	13,333
185	motorbike	0,09110	0,895	45,396	14,000
186	motorcycle	0,01249	0,171	45,567	2,667
187	motorway	0,01361	0,171	45,737	2,667
188	mountain	0,01431	0,128	45,865	2,000
189	museum	0,13699	1,492	47,357	23,333
190	music	0,00325	0,043	47,400	0,667
191	musical	0,00802	0,085	47,485	1,333
192	musical comedy	0,00211	0,043	47,528	0,667

193	near	0,00577	0,043	47,570	0,667
194	neighbour	0,01587	0,171	47,741	2,667
195	neighbourhood	0,05150	0,469	48,210	7,333
196	New York	0,01377	0,171	48,380	2,667
197	newspaper	0,00454	0,085	48,465	1,333
198	nightclub	0,00768	0,085	48,551	1,333
199	noise	0,06808	0,682	49,233	10,667
200	obelisk	0,00489	0,085	49,318	1,333
201	office	0,03711	0,384	49,702	6,000
202	old	0,00281	0,043	49,744	0,667
203	old town	0,00170	0,043	49,787	0,667
204	opera house	0,00262	0,043	49,829	0,667
205	outskirts	0,01143	0,128	49,957	2,000
206	Oxford	0,00667	0,043	50,000	0,667
207	Paris	0,00158	0,043	50,043	0,667
208	park	0,29681	3,069	53,112	48,000
209	parking	0,06691	0,767	53,879	12,000
210	parliament building	0,00281	0,043	53,922	0,667
211	party	0,00325	0,043	53,964	0,667
212	path	0,00537	0,043	54,007	0,667
213	pavement	0,04107	0,384	54,390	6,000
214	pedestrian	0,05325	0,639	55,030	10,000
215	pedestrian area	0,00196	0,043	55,072	0,667
216	pedestrian crossing	0,00841	0,085	55,158	1,333
217	penthouse	0,00325	0,043	55,200	0,667
218	people	0,27362	2,856	58,056	44,667
219	person	0,01386	0,128	58,184	2,000
220	petrol station	0,00281	0,043	58,227	0,667
221	phone box	0,00455	0,085	58,312	1,333
222	place	0,02445	0,171	58,483	2,667
223	plane	0,01121	0,085	58,568	1,333
224	plant	0,00828	0,128	58,696	2,000
225	playground	0,00849	0,085	58,781	1,333
226	pólice	0,05961	0,682	59,463	10,667
227	police station	0,02957	0,384	59,847	6,000
228	policeman	0,02486	0,341	60,188	5,333
229	pollution	0,08379	0,980	61,168	15,333
230	port	0,02826	0,384	61,552	6,000
231	post	0,00244	0,043	61,594	0,667
232	post office	0,03011	0,298	61,893	4,667
233	postbox	0,00325	0,043	61,935	0,667
234	poverty	0,00281	0,043	61,978	0,667
235	problema	0,00684	0,085	62,063	1,333
236	promenade	0,02104	0,213	62,276	3,333
237	pub	0,07992	0,895	63,171	14,000
238	public transport	0,00392	0,085	63,257	1,333
239	public works	0,00211	0,043	63,299	0,667
240	quarter	0,00465	0,043	63,342	0,667
241	radar	0,00433	0,043	63,384	0,667

242	radio	0,00577	0,043	63,427	0,667
243	railway	0,00158	0,043	63,470	0,667
244	railway station	0,00262	0,043	63,512	0,667
245	restaurant	0,15509	1,705	65,217	26,667
246	river	0,02147	0,298	65,516	4,667
247	road	0,21935	2,003	67,519	31,333
248	road sign	0,00227	0,043	67,562	0,667
249	roof	0,00577	0,043	67,604	0,667
250	round	0,00667	0,043	67,647	0,667
251	roundabout	0,04563	0,426	68,073	6,667
252	rubbish	0,01892	0,213	68,286	3,333
253	rush hour	0,00302	0,043	68,329	0,667
254	school	0,13856	1,577	69,906	24,667
255	scooter	0,00325	0,043	69,949	0,667
256	sea	0,01057	0,128	70,077	2,000
257	seafront	0,00808	0,085	70,162	1,333
258	seaside	0,00403	0,043	70,205	0,667
259	secondary school	0,00349	0,043	70,247	0,667
260	semaphore	0,00965	0,085	70,332	1,333
261	semi-detached house	0,00227	0,043	70,375	0,667
262	sewage	0,00325	0,043	70,418	0,667
263	sex	0,00667	0,043	70,460	0,667
264	sex shop	0,00147	0,043	70,503	0,667
265	ship	0,00302	0,043	70,546	0,667
266	shop	0,26184	2,813	73,359	44,000
267	shopping	0,00558	0,085	73,444	1,333
268	shopping centre	0,09356	1,108	74,552	17,333
269	shopping mall	0,00918	0,128	74,680	2,000
270	show	0,00465	0,043	74,723	0,667
271	sidewalk	0,00529	0,085	74,808	1,333
272	sightsee	0,00262	0,043	74,851	0,667
273	sightseeing	0,00667	0,043	74,893	0,667
274	signal	0,01837	0,213	75,107	3,333
275	sky	0,01198	0,085	75,192	1,333
276	skyline	0,00927	0,085	75,277	1,333
277	skyscraper	0,06940	0,639	75,916	10,000
278	smell	0,00158	0,043	75,959	0,667
279	smog	0,00577	0,043	76,002	0,667
280	smoke	0,00577	0,043	76,044	0,667
281	snack bar	0,00433	0,043	76,087	0,667
282	speed	0,00465	0,043	76,130	0,667
283	spring	0,00137	0,043	76,172	0,667
284	square	0,16913	1,662	77,835	26,000
285	stadium	0,00768	0,085	77,920	1,333
286	stair	0,00471	0,085	78,005	1,333

287	stall	0,00183	0,043	78,048	0,667
288	station	0,01143	0,128	78,176	2,000
289	statue	0,02398	0,256	78,431	4,000
290	stop	0,00970	0,085	78,517	1,333
291	store	0,02031	0,213	78,730	3,333
292	Street	0,50812	4,177	82,907	65,333
293	street lamp	0,00349	0,043	82,950	0,667
294	street market	0,00302	0,043	82,992	0,667
295	stress	0,01848	0,213	83,205	3,333
296	stressed people	0,00227	0,043	83,248	0,667
297	suburb	0,01239	0,213	83,461	3,333
298	subway	0,02528	0,298	83,760	4,667
299	supermarket	0,13360	1,577	85,337	24,667
300	swimming pool	0,01350	0,171	85,507	2,667
301	taxi	0,10067	1,194	86,701	18,667
302	taxicab	0,00227	0,043	86,743	0,667
303	tea	0,00375	0,043	86,786	0,667
304	tea room	0,00262	0,043	86,829	0,667
305	the Big Apple	0,00211	0,043	86,871	0,667
306	theatre	0,08494	1,023	87,894	16,000
307	Torremolinos	0,00667	0,043	87,937	0,667
308	tourist	0,00325	0,043	87,980	0,667
309	tower	0,00678	0,085	88,065	1,333
310	town	0,01072	0,128	88,193	2,000
311	town centre	0,00433	0,043	88,235	0,667
312	town hall	0,06724	0,682	88,917	10,667
313	traffic	0,09024	0,980	89,898	15,333
314	traffic jam	0,08003	0,895	90,793	14,000
315	traffic light	0,17243	1,833	92,626	28,667
316	train	0,08700	0,895	93,521	14,000
317	train station	0,06743	0,853	94,373	13,333
318	train stop	0,00620	0,043	94,416	0,667
319	tram	0,00500	0,043	94,459	0,667
320	tramway	0,00262	0,043	94,501	0,667
321	travel	0,00620	0,043	94,544	0,667
322	tree	0,14805	1,620	96,164	25,333
323	truck	0,00500	0,043	96,206	0,667
324	tube	0,03173	0,341	96,547	5,333
325	tube station	0,00500	0,043	96,590	0,667
326	tunnel	0,00196	0,043	96,633	0,667
327	underground	0,11573	1,108	97,741	17,333
328	underground works	0,00158	0,043	97,783	0,667
329	unemployed people	0,00325	0,043	97,826	0,667
330	university	0,09041	1,194	99,020	18,667
331	van	0,00902	0,085	99,105	1,333
332	vehicle	0,00403	0,043	99,147	0,667

333	villa	0,00227	0,043	99,190	0,667
334	village	0,00863	0,128	99,318	2,000
335	wagon	0,00302	0,043	99,361	0,667
336	walking	0,00349	0,043	99,403	0,667
337	way	0,00325	0,043	99,446	0,667
338	Wheel	0,00349	0,043	99,488	0,667
339	window	0,00875	0,085	99,574	1,333
340	woman	0,00244	0,043	99,616	0,667
341	work	0,00603	0,085	99,702	1,333
342	zebra crossing	0,01298	0,171	99,872	2,667
343	zoo	0,00636	0,128	100,000	2,000

2.1.6. Means of transport

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	aeroplane	0,05746	0,716	0,716	8,000
2	aircraft	0,01037	0,119	0,836	1,333
3	airplane	0,14177	1,791	2,627	20,000
4	airport	0,01784	0,299	2,925	3,333
5	ambulance	0,01400	0,239	3,164	2,667
6	animal	0,00697	0,119	3,284	1,333
7	balloon	0,01790	0,299	3,582	3,333
8	bicycle	0,30400	3,701	7,284	41,333
9	bike	0,39361	4,776	12,060	53,333
10	BMW	0,00262	0,060	12,119	0,667
11	board	0,00281	0,060	12,179	0,667
12	boat	0,34843	4,896	17,075	54,667
13	boat train	0,00500	0,060	17,134	0,667
14	bridge	0,00325	0,060	17,194	0,667
15	bulldozer	0,00375	0,060	17,254	0,667
16	bus	0,75300	7,761	25,015	86,667
17	bus station	0,00349	0,060	25,075	0,667
18	bus stop	0,00606	0,119	25,194	1,333
19	cab	0,05939	0,836	26,030	9,333
20	cable car	0,00262	0,060	26,090	0,667
21	caboose	0,00433	0,060	26,149	0,667
22	camel	0,02585	0,358	26,507	4,000
23	camping van	0,00433	0,060	26,567	0,667
24	canoe	0,01683	0,299	26,866	3,333
25	car	0,86363	8,716	35,582	97,333
26	caravan	0,01544	0,239	35,821	2,667
27	carriage	0,00606	0,119	35,940	1,333
28	cart	0,00302	0,060	36,000	0,667
29	chariot	0,00325	0,060	36,060	0,667
30	city	0,00183	0,060	36,119	0,667
31	coach	0,02019	0,299	36,418	3,333
32	cruise	0,02637	0,478	36,896	5,333
33	cruiser	0,03846	0,597	37,493	6,667
34	cycle	0,03909	0,478	37,970	5,333
35	cyclist	0,00620	0,060	38,030	0,667
36	donkey	0,05662	0,896	38,925	10,000
37	double-decker bus	0,01037	0,119	39,045	1,333
38	driver	0,00349	0,060	39,104	0,667
39	dromedary	0,00325	0,060	39,164	0,667
40	duty	0,00262	0,060	39,224	0,667
41	elephant	0,01111	0,179	39,403	2,000
42	elevator	0,00302	0,060	39,463	0,667
43	fare	0,00211	0,060	39,522	0,667
44	fast train	0,01293	0,239	39,761	2,667
45	ferry	0,04601	0,776	40,537	8,667

46	flight	0,00500	0,060	40,597	0,667
47	fly	0,01003	0,119	40,716	1,333
48	go home	0,00302	0,060	40,776	0,667
49	gondola	0,00465	0,060	40,836	0,667
50	harvester	0,00403	0,060	40,896	0,667
51	helicopter	0,16129	2,567	43,463	28,667
52	high-speed train	0,00933	0,119	43,582	1,333
53	highway	0,00349	0,060	43,642	0,667
54	hitchhike	0,00433	0,060	43,701	0,667
55	horse	0,14798	2,269	45,970	25,333
56	horse-drawn carriage	0,00903	0,119	46,090	1,333
57	hovercraft	0,00403	0,060	46,149	0,667
58	hurry	0,00537	0,060	46,209	0,667
59	InterRail	0,00302	0,060	46,269	0,667
60	jeep	0,00433	0,060	46,328	0,667
61	jet	0,03241	0,478	46,806	5,333
62	jet airliner	0,00244	0,060	46,866	0,667
63	jet ski	0,00651	0,119	46,985	1,333
64	jogging	0,00325	0,060	47,045	0,667
65	journey	0,00433	0,060	47,104	0,667
66	kayak	0,00630	0,119	47,224	1,333
67	key	0,00227	0,060	47,284	0,667
68	landing	0,00302	0,060	47,343	0,667
69	leg	0,01086	0,119	47,463	1,333
70	lifeboat	0,00620	0,060	47,522	0,667
71	lift	0,00281	0,060	47,582	0,667
72	limousine	0,00586	0,119	47,701	1,333
73	lorry	0,08143	1,134	48,836	12,667
74	luxury transport	0,00281	0,060	48,896	0,667
75	metro	0,12108	1,552	50,448	17,333
76	metropolitan	0,00375	0,060	50,507	0,667
77	monocycle	0,00158	0,060	50,567	0,667
78	monorail	0,00465	0,060	50,627	0,667
79	moped	0,00577	0,060	50,687	0,667
80	motor	0,00211	0,060	50,746	0,667
81	motor launch	0,00433	0,060	50,806	0,667
82	motorbike	0,40085	5,134	55,940	57,333
83	motorcycle	0,11398	1,493	57,433	16,667
84	motorhome	0,00403	0,060	57,493	0,667
85	motorway	0,00403	0,060	57,552	0,667
86	mountain bike	0,01335	0,179	57,731	2,000
87	move	0,00577	0,060	57,791	0,667
88	ocean liner	0,00933	0,119	57,910	1,333
89	oil	0,00211	0,060	57,970	0,667
90	oil station	0,00196	0,060	58,030	0,667
91	on foot	0,13473	1,970	60,000	22,000
92	parachute	0,01580	0,358	60,358	4,000
93	paraglider	0,00705	0,119	60,478	1,333

94	penny farthing	0,00375	0,060	60,537	0,667
95	plane	0,49180	5,970	66,507	66,667
96	police car	0,00349	0,060	66,567	0,667
97	pony	0,00403	0,060	66,627	0,667
98	private transport	0,00302	0,060	66,687	0,667
99	public transport	0,00325	0,060	66,746	0,667
100	queue	0,00244	0,060	66,806	0,667
101	racing yacht	0,00375	0,060	66,866	0,667
102	railway	0,03342	0,418	67,284	4,667
103	rental car	0,00500	0,060	67,343	0,667
104	road	0,00637	0,119	67,463	1,333
105	rocket	0,01332	0,179	67,642	2,000
106	roller skate	0,02825	0,478	68,119	5,333
107	Rollerblade	0,00684	0,119	68,239	1,333
108	sailing boat	0,00403	0,060	68,299	0,667
109	scooter	0,02020	0,299	68,597	3,333
110	scuba diving	0,00281	0,060	68,657	0,667
111	seat	0,00349	0,060	68,716	0,667
112	seaway	0,00196	0,060	68,776	0,667
113	sheep	0,00620	0,060	68,836	0,667
114	ship	0,26889	3,761	72,597	42,000
115	shipment	0,00325	0,060	72,657	0,667
116	shoe	0,00918	0,179	72,836	2,000
117	shuttle	0,00839	0,119	72,955	1,333
118	sidecar	0,00227	0,060	73,015	0,667
119	skate	0,09685	1,552	74,567	17,333
120	skateboard	0,03063	0,418	74,985	4,667
121	ski	0,01745	0,299	75,284	3,333
122	ski lift	0,00281	0,060	75,343	0,667
123	sky	0,00637	0,119	75,463	1,333
124	sledge	0,00577	0,060	75,522	0,667
125	space shuttle	0,00433	0,060	75,582	0,667
126	spaceship	0,01293	0,239	75,821	2,667
127	stair lift	0,00262	0,060	75,881	0,667
128	submarine	0,03582	0,537	76,418	6,000
129	suburban rail	0,00325	0,060	76,478	0,667
130	subway	0,05305	0,716	77,194	8,000
131	tank	0,00349	0,060	77,254	0,667
132	taxi	0,31296	3,940	81,194	44,000
133	teleport	0,00244	0,060	81,254	0,667
134	ticket	0,00660	0,119	81,373	1,333
135	time machine	0,00375	0,060	81,433	0,667
136	Titanic	0,00158	0,060	81,493	0,667
137	tractor	0,00433	0,060	81,552	0,667
138	traffic lights	0,00244	0,060	81,612	0,667
139	tráiler	0,00875	0,119	81,731	1,333
140	train	0,65866	7,224	88,955	80,667
141	train station	0,00977	0,179	89,134	2,000
142	tram	0,08783	1,313	90,448	14,667

143	tramway	0,00750	0,119	90,567	1,333
144	transatlantic	0,02125	0,299	90,866	3,333
145	Trans-Siberian Railway	0,00325	0,060	90,925	0,667
146	travel	0,00465	0,060	90,985	0,667
147	tricycle	0,01737	0,239	91,224	2,667
148	trike	0,00500	0,060	91,284	0,667
149	trip	0,00403	0,060	91,343	0,667
150	trolley	0,01657	0,239	91,582	2,667
151	trolleybus	0,00375	0,060	91,642	0,667
152	truck	0,07550	1,254	92,896	14,000
153	tube	0,11093	1,493	94,388	16,667
154	underground	0,25290	3,224	97,612	36,000
155	underway	0,00403	0,060	97,672	0,667
156	van	0,03451	0,537	98,209	6,000
157	vehicle	0,00537	0,060	98,269	0,667
158	Vespa	0,00302	0,060	98,328	0,667
159	wagon	0,00302	0,060	98,388	0,667
160	walk	0,00841	0,119	98,507	1,333
161	walking frame	0,01150	0,179	98,687	2,000
162	water	0,00147	0,060	98,746	0,667
163	wheelchair	0,00262	0,060	98,806	0,667
164	yacht	0,07141	1,134	99,940	12,667
165	Zeppelin	0,00433	0,060	100,000	0,667

2.1.7. Games and entertainments

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	aerobics	0,03835	0,523	6,623	60,234
2	aerodance	0,00225	0,052	0,662	54,386
3	American football	0,01031	0,105	1,325	49,708
4	amusement park	0,00242	0,052	0,662	50,292
5	appointment with friends	0,00260	0,052	0,662	40,936
6	arcade	0,00195	0,052	0,662	36,842
7	art	0,00869	0,105	1,325	35,088
8	art gallery	0,00373	0,052	0,662	35,673
9	athletics	0,00704	0,105	1,325	32,749
10	backgammon	0,02584	0,261	3,311	32,164
11	badminton	0,01654	0,209	2,649	32,749
12	ball	0,08918	0,889	11,258	30,409
13	ballet	0,01002	0,157	1,987	29,825
14	bar	0,01964	0,314	3,974	27,485
15	baseball	0,09195	1,045	13,245	24,561
16	basketball	0,43288	4,443	56,291	23,392
17	beach	0,02107	0,314	3,974	21,637
18	bicycle	0,00534	0,052	0,662	19,298
19	bike	0,02149	0,209	2,649	20,468
20	billiard	0,00945	0,157	1,987	20,468
21	bingo	0,00347	0,052	0,662	16,959
22	blackjack	0,00813	0,105	1,325	18,713
23	blind man's buff	0,00797	0,105	1,325	17,544
24	blog	0,00462	0,052	0,662	18,713
25	board game	0,01036	0,105	1,325	15,205
26	book	0,03774	0,366	4,636	18,713
27	bowling	0,03471	0,470	5,960	17,544
28	boxing	0,00462	0,052	0,662	15,789
29	Breach	0,00242	0,052	0,662	15,789
30	bungee jumping	0,00462	0,052	0,662	14,035
31	Call of Duty	0,00662	0,052	0,662	12,865
32	camping	0,00776	0,105	1,325	13,450
33	car	0,00300	0,052	0,662	11,696
34	card	0,17485	1,934	24,503	11,111
35	card game	0,01147	0,105	1,325	12,281
36	catch	0,00574	0,052	0,662	12,281
37	CD	0,00260	0,052	0,662	10,526
38	charade	0,00323	0,052	0,662	10,526
39	chat	0,02071	0,314	3,974	10,526
40	chatter	0,00612	0,105	1,325	11,696
41	chess	0,18315	1,725	21,854	11,111
42	children	0,00759	0,105	1,325	9,942
43	Chinese food	0,00195	0,052	0,662	11,111

44	cinema	0,33948	3,868	49,007	10,526
45	circus	0,02269	0,261	3,311	10,526
46	clean	0,00280	0,052	0,662	10,526
47	climbing	0,00809	0,105	1,325	9,357
48	club	0,00210	0,052	0,662	9,357
49	clubbing	0,00323	0,052	0,662	8,772
50	Cluedo	0,00430	0,052	0,662	8,772
51	coach	0,00300	0,052	0,662	8,772
52	comic	0,00225	0,052	0,662	7,602
53	competition	0,00157	0,052	0,662	7,602
54	computer	0,07475	0,836	10,596	9,357
55	computer game	0,08028	0,836	10,596	8,187
56	Conan	0,00616	0,052	0,662	8,187
57	concert	0,08846	1,150	14,570	8,187
58	console	0,00616	0,052	0,662	6,433
59	cooking	0,02297	0,314	3,974	7,602
60	Crash Team Racing	0,00534	0,052	0,662	7,018
61	crawl	0,00323	0,052	0,662	7,018
62	cricket	0,01208	0,157	1,987	7,602
63	crossword	0,02934	0,366	4,636	7,018
64	cycling	0,05171	0,732	9,272	6,433
65	dance	0,14827	2,143	27,152	7,018
66	dart	0,00400	0,052	0,662	7,018
67	Desperate Housewives	0,00225	0,052	0,662	5,263
68	dice	0,03043	0,366	4,636	7,602
69	digital booklet	0,00574	0,052	0,662	5,848
70	disco(theque)	0,04015	0,575	7,285	6,433
71	diving	0,00547	0,105	1,325	6,433
72	doll	0,03436	0,418	5,298	5,263
73	dominoes	0,01898	0,157	1,987	7,018
74	drawing	0,03115	0,418	5,298	4,678
75	drink	0,01005	0,157	1,987	4,678
76	drinking bout	0,00260	0,052	0,662	5,263
77	drinking game	0,00672	0,105	1,325	5,263
78	drive	0,00430	0,052	0,662	5,263
79	drum	0,00996	0,105	1,325	4,678
80	DVD	0,00896	0,105	1,325	5,848
81	eat	0,00300	0,052	0,662	5,263
82	eat out	0,00640	0,105	1,325	4,678
83	exercise	0,00260	0,052	0,662	5,263
84	extreme sport	0,00300	0,052	0,662	5,263
85	Facebook	0,02610	0,366	4,636	5,263
86	fashion designer	0,00400	0,052	0,662	5,263
87	festival	0,00972	0,157	1,987	5,848
88	field	0,00400	0,052	0,662	5,263
89	fight	0,00242	0,052	0,662	4,678
90	film	0,07326	0,784	9,934	4,678
91	fishing	0,01479	0,261	3,311	4,094

92	fitness	0,00940	0,157	1,987	3,509
93	football	0,55680	5,332	67,550	4,094
94	football league	0,00323	0,052	0,662	4,094
95	football match	0,01297	0,157	1,987	4,094
96	Formula 1	0,00534	0,052	0,662	3,509
97	friend	0,00869	0,105	1,325	4,678
98	fun run	0,00210	0,052	0,662	3,509
99	game	0,01190	0,105	1,325	3,509
100	Game Boy	0,00373	0,052	0,662	2,924
101	gin rummy	0,00400	0,052	0,662	4,094
102	go for a walk	0,01160	0,157	1,987	4,094
103	go out	0,04030	0,523	6,623	4,094
104	go out for a drink	0,00260	0,052	0,662	3,509
105	go out to dinner	0,00785	0,157	1,987	3,509
106	go out with friends	0,03895	0,680	8,609	3,509
107	go shopping	0,03253	0,418	5,298	2,924
108	golf	0,05292	0,627	7,947	3,509
109	Greek food	0,00182	0,052	0,662	2,924
110	Green Farm	0,00462	0,052	0,662	3,509
111	Guess Who?	0,00373	0,052	0,662	3,509
112	guitar	0,00844	0,105	1,325	3,509
113	Guitar Hero	0,00373	0,052	0,662	2,924
114	gym	0,02104	0,418	5,298	2,924
115	gymnastics	0,00777	0,105	1,325	2,924
116	handball	0,03398	0,418	5,298	2,924
117	hang out with somebody	0,00280	0,052	0,662	2,924
118	hide	0,00347	0,052	0,662	2,339
119	hide-and-seek	0,08412	0,889	11,258	2,339
120	hiking	0,01264	0,157	1,987	2,924
121	hockey	0,04402	0,523	6,623	2,924
122	horse	0,00462	0,052	0,662	2,924
123	horse racing	0,00497	0,052	0,662	2,339
124	horse riding	0,00896	0,105	1,325	2,924
125	How I Met Your Mother	0,00462	0,052	0,662	2,339
126	hunting	0,00347	0,052	0,662	2,924
127	ice hockey	0,00347	0,052	0,662	2,339
128	ice skating	0,01501	0,157	1,987	2,924
129	Internet	0,07342	0,993	12,583	2,924
130	jazz concert	0,00462	0,052	0,662	2,924
131	jogging	0,10001	1,411	17,881	2,339
132	joke	0,00662	0,052	0,662	2,924
133	jumping	0,00505	0,105	1,325	2,924
134	Just Dance	0,00195	0,052	0,662	2,339
135	karaoke	0,04775	0,627	7,947	2,339
136	karting	0,00776	0,105	1,325	2,924
137	kayak	0,00657	0,105	1,325	2,924
138	keyboard	0,00574	0,052	0,662	2,339

139	lake	0,00323	0,052	0,662	2,924
140	laptop	0,00497	0,052	0,662	2,339
141	Lost	0,00534	0,052	0,662	2,339
142	ludo	0,02567	0,261	3,311	1,754
143	magazine	0,00347	0,052	0,662	2,339
144	match	0,00400	0,052	0,662	2,339
145	meeting	0,00323	0,052	0,662	2,339
146	merry-go-round	0,00662	0,052	0,662	2,339
147	Monkey Island	0,00616	0,052	0,662	2,924
148	Monopoly	0,05999	0,732	9,272	1,754
149	motor racing	0,00574	0,052	0,662	2,339
150	motorcycling	0,00985	0,105	1,325	2,339
151	mountain	0,00799	0,105	1,325	2,339
152	movie	0,06700	0,784	9,934	1,754
153	museum	0,01500	0,261	3,311	2,339
154	music	0,25773	3,136	39,735	2,339
155	musical comedy	0,01089	0,157	1,987	1,754
156	nature	0,00430	0,052	0,662	2,339
157	network	0,00462	0,052	0,662	1,754
158	newspaper	0,00373	0,052	0,662	1,754
159	nightclub	0,01107	0,105	1,325	1,754
160	Nintendo	0,00820	0,105	1,325	1,754
161	opera	0,03948	0,575	7,285	1,754
162	paddle	0,07405	0,889	11,258	2,339
163	paint	0,00373	0,052	0,662	1,754
164	painting	0,03731	0,418	5,298	1,754
165	paragliding	0,00373	0,052	0,662	1,754
166	Parcheesi	0,02340	0,209	2,649	1,754
167	park	0,00858	0,157	1,987	1,754
168	party	0,01938	0,209	2,649	2,339
169	password	0,00528	0,105	1,325	1,754
170	personal computer	0,00462	0,052	0,662	1,754
171	pétanque	0,01555	0,157	1,987	1,754
172	Petz	0,00430	0,052	0,662	2,339
173	phone	0,00260	0,052	0,662	1,754
174	photography	0,01718	0,209	2,649	1,754
175	piano	0,00534	0,052	0,662	1,754
176	Pictionary	0,00534	0,052	0,662	1,754
177	Pilates	0,00373	0,052	0,662	1,754
178	pinball	0,00497	0,052	0,662	1,754
179	ping-pong	0,01231	0,157	1,987	1,170
180	play	0,02442	0,261	3,311	1,754
181	play a musical instrument	0,02676	0,314	3,974	1,754
182	play music	0,01532	0,209	2,649	1,170
183	Play Station	0,13897	1,464	18,543	1,754
184	play the cello	0,00430	0,052	0,662	1,754
185	play the clarinet	0,00242	0,052	0,662	1,170
186	play the flute	0,00400	0,052	0,662	1,754

187	play the guitar	0,02381	0,261	3,311	1,754
188	play the piano	0,01279	0,209	2,649	1,754
189	playground	0,00574	0,052	0,662	1,754
190	Playmobil	0,00347	0,052	0,662	2,339
191	poker	0,07855	0,836	10,596	1,170
192	polo	0,01477	0,157	1,987	1,754
193	pub	0,03907	0,523	6,623	1,170
194	puzzle	0,01422	0,209	2,649	1,754
195	quest	0,00497	0,052	0,662	1,754
196	Rabbids	0,00497	0,052	0,662	1,170
197	racing	0,00497	0,052	0,662	1,754
198	racket	0,00964	0,105	1,325	1,170
199	radio	0,05367	0,836	10,596	1,170
200	rafting	0,00462	0,052	0,662	1,170
201	rally	0,00731	0,105	1,325	1,754
202	read	0,00532	0,105	1,325	1,170
203	read a book	0,03406	0,418	5,298	1,170
204	reading	0,13871	1,673	21,192	1,170
205	relax	0,00574	0,052	0,662	1,170
206	restaurant	0,01582	0,209	2,649	1,170
207	ride a horse	0,01644	0,261	3,311	1,170
208	riding	0,01086	0,157	1,987	1,170
209	Risk	0,00400	0,052	0,662	1,170
210	role play	0,02802	0,261	3,311	1,754
211	roller skating	0,00863	0,105	1,325	1,170
212	Rollerblade	0,00906	0,105	1,325	1,170
213	roulette	0,00462	0,052	0,662	1,170
214	rugby	0,06212	0,732	9,272	1,170
215	rugby match	0,00534	0,052	0,662	1,170
216	running	0,10046	1,307	16,556	1,170
217	sailing	0,01002	0,157	1,987	1,170
218	sandbox	0,00400	0,052	0,662	1,170
219	Scrabble	0,02388	0,261	3,311	1,170
220	scuba diving	0,00430	0,052	0,662	1,170
221	sculpture	0,00300	0,052	0,662	1,754
222	sea	0,00616	0,052	0,662	1,170
223	seesaw	0,00616	0,052	0,662	1,170
224	serial	0,00574	0,052	0,662	1,754
225	series	0,02332	0,261	3,311	1,170
226	shop	0,00323	0,052	0,662	1,170
227	shopping centre	0,00300	0,052	0,662	1,170
228	show	0,01413	0,157	1,987	1,170
229	sightseeing	0,00580	0,105	1,325	1,170
230	Simon says	0,00863	0,105	1,325	1,170
231	Sims	0,00868	0,157	1,987	1,170
232	sing	0,02072	0,261	3,311	1,170
233	singing	0,06905	0,941	11,921	1,754
234	skate	0,03849	0,523	6,623	1,170
235	skating	0,01431	0,157	1,987	1,170

236	ski	0,01715	0,261	3,311	1,170
237	skiing	0,03843	0,470	5,960	1,170
238	Skype	0,01295	0,157	1,987	1,170
239	slalom	0,00497	0,052	0,662	1,170
240	sleep	0,00400	0,052	0,662	1,754
241	snakes and ladders	0,00497	0,052	0,662	1,170
242	snorkel	0,00136	0,052	0,662	1,170
243	snow	0,00534	0,052	0,662	1,170
244	snowboard	0,01586	0,209	2,649	1,170
245	soap opera	0,01538	0,157	1,987	1,170
246	soccer	0,06921	0,732	9,272	1,170
247	social network	0,00497	0,052	0,662	1,170
248	softball	0,00939	0,105	1,325	1,170
249	speak	0,00784	0,157	1,987	1,170
250	speak with friends	0,00260	0,052	0,662	1,170
251	Spinning	0,02257	0,366	4,636	0,585
252	sport	0,18367	2,091	26,490	0,585
253	squash	0,00673	0,105	1,325	0,585
254	Starbucks	0,00323	0,052	0,662	1,170
255	step	0,00843	0,157	1,987	1,170
256	strategy game	0,00497	0,052	0,662	0,585
257	study	0,00242	0,052	0,662	0,585
258	sudoku	0,01347	0,157	1,987	0,585
259	sunbathing	0,00242	0,052	0,662	0,585
260	Super Mario Bros	0,01157	0,157	1,987	0,585
261	surf	0,01477	0,209	2,649	0,585
262	surf the Net	0,00831	0,105	1,325	0,585
263	surfing	0,01192	0,157	1,987	0,585
264	swim	0,01159	0,105	1,325	1,170
265	swimming	0,12406	1,673	21,192	1,170
266	swimming pool	0,01273	0,157	1,987	0,585
267	swing	0,00574	0,052	0,662	0,585
268	table football	0,00347	0,052	0,662	0,585
269	table game	0,02970	0,314	3,974	1,170
270	table tennis	0,01959	0,261	3,311	1,170
271	Taboo	0,02726	0,261	3,311	1,170
272	take a coffee	0,00347	0,052	0,662	1,170
273	take a nap	0,00347	0,052	0,662	0,585
274	take a picture	0,00400	0,052	0,662	0,585
275	talk	0,00598	0,105	1,325	0,585
276	talk with friends	0,00623	0,105	1,325	0,585
277	team	0,00280	0,052	0,662	0,585
278	television	0,24257	3,032	38,411	0,585
279	television broadcast	0,00616	0,052	0,662	0,585
280	television programme	0,00574	0,052	0,662	0,585
281	tennis	0,33085	3,607	45,695	0,585

282	Texas hold'em	0,00260	0,052	0,662	0,585
283	The Legend of Kyrandia	0,00662	0,052	0,662	0,585
284	theatre	0,18061	2,196	27,815	0,585
285	trading card game	0,00497	0,052	0,662	0,585
286	training	0,00462	0,052	0,662	0,585
287	travel	0,05407	0,732	9,272	0,585
288	trekking	0,01530	0,209	2,649	0,585
289	triathlon	0,00300	0,052	0,662	0,585
290	trick or treat	0,00534	0,052	0,662	1,170
291	Trivial	0,03052	0,314	3,974	1,170
292	Tuenti	0,01050	0,157	1,987	0,585
293	Twitter	0,00430	0,052	0,662	0,585
294	video	0,01539	0,157	1,987	0,585
295	videogame	0,30921	3,032	38,411	0,585
296	volleyball	0,12142	1,411	17,881	0,585
297	walk	0,00347	0,052	0,662	0,585
298	walking	0,05291	0,680	8,609	0,585
299	water polo	0,01850	0,209	2,649	0,585
300	Wii	0,05334	0,627	7,947	0,585
301	window-shop	0,00323	0,052	0,662	1,170
302	windsurf	0,00869	0,105	1,325	0,585
303	wine and dine	0,00146	0,052	0,662	0,585
304	World of Warcraft	0,01159	0,105	1,325	0,585
305	wrestling	0,00225	0,052	0,662	0,585
306	write a book	0,00462	0,052	0,662	0,585
307	writing	0,03016	0,470	5,960	0,585
308	Xbox	0,02324	0,261	3,311	0,585
309	Yahtzee	0,00430	0,052	0,662	0,585
310	yoga	0,01278	0,209	2,649	0,585

2.1.8. Professions and jobs

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	accountant	0,01223	0,155	0,155	2,000
2	actor	0,15542	2,073	2,228	26,667
3	actress	0,02897	0,415	2,642	5,333
4	administrative assistant	0,00262	0,052	2,694	0,667
5	administrator	0,00302	0,052	2,746	0,667
6	advertiser	0,00620	0,052	2,798	0,667
7	adviser	0,00262	0,052	2,850	0,667
8	agent	0,00577	0,052	2,902	0,667
9	allergist	0,00325	0,052	2,953	0,667
10	anchorman	0,00244	0,052	3,005	0,667
11	annalist	0,00500	0,052	3,057	0,667
12	architect	0,17297	1,917	4,974	24,667
13	artista	0,07375	0,984	5,959	12,667
14	assessor	0,00465	0,052	6,010	0,667
15	assistant	0,04394	0,518	6,528	6,667
16	astronaut	0,00211	0,052	6,580	0,667
17	athlete	0,00196	0,052	6,632	0,667
18	attendant	0,00227	0,052	6,684	0,667
19	attorney	0,00500	0,052	6,736	0,667
20	au pair	0,00605	0,104	6,839	1,333
21	babysitter	0,06110	0,881	7,720	11,333
22	Baker	0,07489	0,984	8,705	12,667
23	bank robber	0,00537	0,052	8,756	0,667
24	banker	0,01095	0,259	9,016	3,333
25	barman	0,08148	0,984	10,000	12,667
26	barrister	0,01373	0,155	10,155	2,000
27	bartender	0,01012	0,155	10,311	2,000
28	basketball player	0,01188	0,155	10,466	2,000
29	biologist	0,00244	0,052	10,518	0,667
30	blacksmith	0,00537	0,052	10,570	0,667
31	blue-collar worker	0,00325	0,052	10,622	0,667
32	bookkeeper	0,00433	0,052	10,674	0,667
33	boss	0,02160	0,259	10,933	3,333
34	boxer	0,00880	0,104	11,036	1,333
35	bricklayer	0,00981	0,155	11,192	2,000
36	broadcaster	0,00227	0,052	11,244	0,667
37	broker	0,02170	0,311	11,554	4,000
38	builder	0,01603	0,259	11,813	3,333
39	bullfighter	0,01738	0,259	12,073	3,333
40	bus driver	0,08817	1,036	13,109	13,333
41	businessman	0,03248	0,466	13,575	6,000
42	butcher	0,10108	1,140	14,715	14,667
43	butler	0,00500	0,052	14,767	0,667
44	buyer	0,00244	0,052	14,819	0,667
45	cameraman	0,00403	0,052	14,870	0,667

46	carpenter	0,05050	0,518	15,389	6,667
47	cashier	0,00836	0,104	15,492	1,333
48	cellist	0,00577	0,052	15,544	0,667
49	chauffeur	0,01185	0,155	15,699	2,000
50	chef	0,04842	0,622	16,321	8,000
51	chemical engineer	0,00620	0,052	16,373	0,667
52	chemist	0,02546	0,415	16,788	5,333
53	chief communications officer	0,01654	0,207	16,995	2,667
54	civil servant	0,00931	0,104	17,098	1,333
55	cleaner	0,04757	0,674	17,772	8,667
56	Clerk	0,02799	0,363	18,135	4,667
57	clown	0,00580	0,104	18,238	1,333
58	coach	0,01870	0,311	18,549	4,000
59	coach driver	0,00500	0,052	18,601	0,667
60	comedian	0,00211	0,052	18,653	0,667
61	commercial traveller	0,00537	0,052	18,705	0,667
62	community manager	0,00375	0,052	18,756	0,667
63	computer programmer	0,01425	0,155	18,912	2,000
64	computer scientist	0,00403	0,052	18,964	0,667
65	concierge	0,00651	0,104	19,067	1,333
66	conductor	0,02006	0,259	19,326	3,333
67	construction engineer	0,00667	0,052	19,378	0,667
68	construction worker	0,00433	0,052	19,430	0,667
69	constructor	0,01060	0,155	19,585	2,000
70	consultant	0,00375	0,052	19,637	0,667
71	cook	0,09965	1,295	20,933	16,667
72	costume designer	0,00262	0,052	20,984	0,667
73	counsellor	0,00349	0,052	21,036	0,667
74	curator	0,00700	0,104	21,140	1,333
75	dancer	0,07317	0,933	22,073	12,000
76	decorator	0,00325	0,052	22,124	0,667
77	dentist	0,05686	0,674	22,798	8,667
78	designer	0,07073	1,036	23,834	13,333
79	developer	0,01333	0,104	23,938	1,333
80	director	0,04247	0,570	24,508	7,333
81	doctor	0,52568	5,337	29,845	68,667
82	domestic worker	0,00244	0,052	29,896	0,667
83	doorman	0,00724	0,104	30,000	1,333
84	draughtsman	0,01371	0,207	30,207	2,667
85	driver	0,20009	2,332	32,539	30,000
86	dustman	0,01320	0,155	32,694	2,000
87	ear nose and throat specialist	0,00375	0,052	32,746	0,667
88	economist	0,04707	0,725	33,472	9,333
89	electrical engineer	0,00537	0,052	33,523	0,667

90	electrician	0,04496	0,518	34,041	6,667
91	employee	0,00631	0,104	34,145	1,333
92	employer	0,00620	0,052	34,197	0,667
93	engineer	0,20241	2,228	36,425	28,667
94	entrepreneur	0,00568	0,155	36,580	2,000
95	environmentalist	0,00262	0,052	36,632	0,667
96	explorer	0,00500	0,052	36,684	0,667
97	factory worker	0,00403	0,052	36,736	0,667
98	farmer	0,05254	0,674	37,409	8,667
99	fashion model	0,00325	0,052	37,461	0,667
100	fat cat	0,00375	0,052	37,513	0,667
101	fighter	0,00651	0,104	37,617	1,333
102	fighter pilot	0,00325	0,052	37,668	0,667
103	film critic	0,00227	0,052	37,720	0,667
104	film director	0,00577	0,052	37,772	0,667
105	film producer	0,00537	0,052	37,824	0,667
106	film star	0,00262	0,052	37,876	0,667
107	film-maker	0,00375	0,052	37,927	0,667
108	financial adviser	0,00500	0,052	37,979	0,667
109	firefighter	0,04147	0,518	38,497	6,667
110	fireman	0,16695	2,021	40,518	26,000
111	fisherman	0,02550	0,311	40,829	4,000
112	fishmonger	0,00933	0,104	40,933	1,333
113	flight attendant	0,01410	0,259	41,192	3,333
114	football player	0,07722	0,881	42,073	11,333
115	footballer	0,04916	0,622	42,694	8,000
116	forensic	0,00500	0,052	42,746	0,667
117	freelancer	0,00536	0,104	42,850	1,333
118	gardener	0,03926	0,518	43,368	6,667
119	geographer	0,01752	0,207	43,575	2,667
120	geologist	0,00170	0,052	43,627	0,667
121	graphic designer	0,01052	0,155	43,782	2,000
122	gravedigger	0,00667	0,052	43,834	0,667
123	greengrocer	0,00349	0,052	43,886	0,667
124	grocer	0,00325	0,052	43,938	0,667
125	guide	0,01618	0,207	44,145	2,667
126	guitarist	0,02674	0,363	44,508	4,667
127	gynaecologist	0,00403	0,052	44,560	0,667
128	hacker	0,00281	0,052	44,611	0,667
129	hairdresser	0,04548	0,518	45,130	6,667
130	headmaster	0,00840	0,104	45,233	1,333
131	headmistress	0,00500	0,052	45,285	0,667
132	historian	0,01961	0,207	45,492	2,667
133	househusband	0,00403	0,052	45,544	0,667
134	housekeeper	0,02074	0,311	45,855	4,000
135	housewife	0,04670	0,622	46,477	8,000
136	iceman	0,00196	0,052	46,528	0,667
137	illustrator	0,00577	0,052	46,580	0,667
138	interpreter	0,03706	0,466	47,047	6,000

139	job holder	0,00403	0,052	47,098	0,667
140	journalist	0,07591	0,984	48,083	12,667
141	judge	0,05012	0,622	48,705	8,000
142	King	0,00433	0,052	48,756	0,667
143	kitchen maid	0,00244	0,052	48,808	0,667
144	labourer	0,00577	0,052	48,860	0,667
145	lawyer	0,25652	2,798	51,658	36,000
146	lecturer	0,00778	0,155	51,813	2,000
147	librarian	0,00403	0,052	51,865	0,667
148	lieutenant	0,00465	0,052	51,917	0,667
149	lifeguard	0,00128	0,052	51,969	0,667
150	lifesaver	0,00227	0,052	52,021	0,667
151	linguist	0,01494	0,207	52,228	2,667
152	locksmith	0,00880	0,104	52,332	1,333
153	lorry driver	0,01315	0,155	52,487	2,000
154	magician	0,00137	0,052	52,539	0,667
155	manager	0,08324	1,036	53,575	13,333
156	marketing manager	0,00403	0,052	53,627	0,667
157	marshal	0,00465	0,052	53,679	0,667
158	masseur	0,00158	0,052	53,731	0,667
159	master	0,00403	0,052	53,782	0,667
160	mathematician	0,01282	0,155	53,938	2,000
161	mayor	0,01859	0,259	54,197	3,333
162	mechanic	0,07804	0,933	55,130	12,000
163	Member of Parliament	0,00465	0,052	55,181	0,667
164	midwife	0,00244	0,052	55,233	0,667
165	military policeman	0,00627	0,104	55,337	1,333
166	minister	0,01127	0,155	55,492	2,000
167	model	0,00433	0,052	55,544	0,667
168	monk	0,00170	0,052	55,596	0,667
169	mother	0,00302	0,052	55,648	0,667
170	musician	0,07540	0,881	56,528	11,333
171	nanny	0,00840	0,104	56,632	1,333
172	neurologist	0,00349	0,052	56,684	0,667
173	news agent	0,00577	0,052	56,736	0,667
174	novelist	0,00433	0,052	56,788	0,667
175	nun	0,01307	0,259	57,047	3,333
176	nurse	0,26060	2,902	59,948	37,333
177	nursery nurse	0,00667	0,052	60,000	0,667
178	nutritionist	0,00500	0,052	60,052	0,667
179	oculist	0,00375	0,052	60,104	0,667
180	office worker	0,00500	0,052	60,155	0,667
181	officer	0,00577	0,052	60,207	0,667
182	official	0,00349	0,052	60,259	0,667
183	paediatrician	0,00302	0,052	60,311	0,667
184	painter	0,11890	1,451	61,762	18,667
185	paramedic	0,00537	0,052	61,813	0,667
186	pensioner	0,00375	0,052	61,865	0,667

187	personal assistant	0,00665	0,104	61,969	1,333
188	personal shopper	0,00525	0,104	62,073	1,333
189	personal trainer	0,00577	0,052	62,124	0,667
190	pharmacist	0,00440	0,104	62,228	1,333
191	philosopher	0,00537	0,052	62,280	0,667
192	photographer	0,01373	0,155	62,435	2,000
193	physician	0,00525	0,104	62,539	1,333
194	physicist	0,01929	0,207	62,746	2,667
195	pianist	0,01549	0,207	62,953	2,667
196	piano player	0,01095	0,155	63,109	2,000
197	pilot	0,07856	0,984	64,093	12,667
198	pimp	0,00620	0,052	64,145	0,667
199	player	0,00403	0,052	64,197	0,667
200	plumber	0,13514	1,399	65,596	18,000
201	poet	0,00733	0,155	65,751	2,000
202	police officer	0,00902	0,104	65,855	1,333
203	policeman	0,37846	4,041	69,896	52,000
204	policewoman	0,00433	0,052	69,948	0,667
205	politician	0,06152	0,829	70,777	10,667
206	porn actress	0,00500	0,052	70,829	0,667
207	porter	0,00227	0,052	70,881	0,667
208	postman	0,03465	0,415	71,295	5,333
209	presenter	0,04306	0,674	71,969	8,667
210	president	0,02296	0,311	72,280	4,000
211	priest	0,00705	0,155	72,435	2,000
212	prime minister	0,00349	0,052	72,487	0,667
213	principal	0,00281	0,052	72,539	0,667
214	producer	0,01118	0,155	72,694	2,000
215	professor	0,13917	1,554	74,249	20,000
216	programmer	0,00537	0,052	74,301	0,667
217	prostitute	0,01503	0,155	74,456	2,000
218	psychiatrist	0,00403	0,052	74,508	0,667
219	psychologist	0,06467	0,725	75,233	9,333
220	public servant	0,00281	0,052	75,285	0,667
221	receptionist	0,03293	0,466	75,751	6,000
222	reporter	0,00674	0,104	75,855	1,333
223	researcher	0,02583	0,311	76,166	4,000
224	robber	0,00577	0,052	76,218	0,667
225	runner	0,00211	0,052	76,269	0,667
226	sailor	0,01287	0,104	76,373	1,333
227	sales manager	0,00728	0,104	76,477	1,333
228	salesman	0,04081	0,466	76,943	6,000
229	salesperson	0,00403	0,052	76,995	0,667
230	scientist	0,02421	0,363	77,358	4,667
231	sculptor	0,00403	0,052	77,409	0,667
232	secretary	0,06764	0,881	78,290	11,333
233	security guard	0,01767	0,207	78,497	2,667
234	seller	0,01732	0,259	78,756	3,333
235	servant	0,00465	0,052	78,808	0,667

236	shoemaker	0,00196	0,052	78,860	0,667
237	shop assistant	0,10092	1,295	80,155	16,667
238	shopper	0,00584	0,104	80,259	1,333
239	showman	0,00465	0,052	80,311	0,667
240	singer	0,11453	1,399	81,710	18,000
241	soccer player	0,00577	0,052	81,762	0,667
242	social worker	0,00433	0,052	81,813	0,667
243	software engineer	0,00667	0,052	81,865	0,667
244	soldier	0,01482	0,259	82,124	3,333
245	solicitor	0,02975	0,415	82,539	5,333
246	soloist	0,00375	0,052	82,591	0,667
247	speaker	0,00620	0,052	82,642	0,667
248	sportsman	0,02855	0,363	83,005	4,667
249	steward	0,01000	0,104	83,109	1,333
250	stripper	0,00537	0,052	83,161	0,667
251	student	0,06085	0,725	83,886	9,333
252	stylist	0,00577	0,052	83,938	0,667
253	supermodel	0,00602	0,104	84,041	1,333
254	surveyor	0,00620	0,052	84,093	0,667
255	systems engineer	0,00537	0,052	84,145	0,667
256	tailor	0,00953	0,104	84,249	1,333
257	tattooist	0,00537	0,052	84,301	0,667
258	taxi driver	0,11802	1,399	85,699	18,000
259	teacher	0,84771	7,409	93,109	95,333
260	technician	0,00465	0,052	93,161	0,667
261	tennis player	0,01694	0,207	93,368	2,667
262	tester	0,00349	0,052	93,420	0,667
263	thief	0,00371	0,104	93,523	1,333
264	tourist agent	0,01302	0,155	93,679	2,000
265	tourist guide	0,01829	0,259	93,938	3,333
266	trainer	0,00700	0,104	94,041	1,333
267	translator	0,06690	0,829	94,870	10,667
268	travel agent	0,01236	0,155	95,026	2,000
269	troubadour	0,00262	0,052	95,078	0,667
270	typist	0,00500	0,052	95,130	0,667
271	veterinarian	0,00700	0,104	95,233	1,333
272	violinist	0,02419	0,259	95,492	3,333
273	waiter	0,18576	2,228	97,720	28,667
274	waitress	0,01816	0,311	98,031	4,000
275	webmaster	0,00620	0,052	98,083	0,667
276	white-collar worker	0,00302	0,052	98,135	0,667
277	witch doctor	0,00375	0,052	98,187	0,667
278	worker	0,01417	0,155	98,342	2,000
279	workman	0,00819	0,104	98,446	1,333
280	writer	0,12553	1,554	100,000	20,000

2.1.9. Computer and internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(Adobe) Photoshop	0,00866	0,089	0,089	1,333
2	(Microsoft) Access	0,01673	0,222	0,311	3,333
3	(Microsoft) Excel	0,01954	0,178	0,488	2,667
4	(Microsoft) Office	0,00875	0,089	0,577	1,333
5	(Microsoft) PowerPoint	0,00869	0,089	0,666	1,333
6	(Microsoft) Word	0,04569	0,488	1,155	7,333
7	account	0,00111	0,044	1,199	0,667
8	Acer	0,00403	0,044	1,243	0,667
9	add	0,00366	0,089	1,332	1,333
10	address	0,00325	0,044	1,377	0,667
11	Adobe	0,00281	0,044	1,421	0,667
12	ADSL	0,00758	0,089	1,510	1,333
13	advertisement	0,00563	0,089	1,599	1,333
14	Apple	0,03235	0,355	1,954	5,333
15	applet	0,00500	0,044	1,998	0,667
16	application	0,01211	0,133	2,131	2,000
17	Ares	0,00128	0,044	2,176	0,667
18	article	0,00325	0,044	2,220	0,667
19	ASUS	0,00375	0,044	2,265	0,667
20	attach	0,00496	0,089	2,353	1,333
21	audio	0,00465	0,044	2,398	0,667
22	auricular	0,00244	0,044	2,442	0,667
23	backup copy	0,00537	0,044	2,487	0,667
24	band	0,00325	0,044	2,531	0,667
25	banner	0,00577	0,044	2,575	0,667
26	battery	0,01016	0,133	2,709	2,000
27	Big Brother	0,00433	0,044	2,753	0,667
28	Bill Gates	0,00584	0,089	2,842	1,333
29	bit	0,02582	0,266	3,108	4,000
30	blog	0,10801	1,066	4,174	16,000
31	blogger	0,02337	0,266	4,440	4,000
32	Blue-ray Disc	0,00724	0,089	4,529	1,333
33	Bluetooth	0,01061	0,133	4,663	2,000
34	board	0,03198	0,311	4,973	4,667
35	book	0,00752	0,089	5,062	1,333
36	bookmark	0,00433	0,044	5,107	0,667
37	box	0,00281	0,044	5,151	0,667
38	broadband	0,00637	0,089	5,240	1,333
39	browse	0,01273	0,133	5,373	2,000
40	browser	0,03266	0,355	5,728	5,333
41	button	0,00375	0,044	5,773	0,667
42	byte	0,01184	0,133	5,906	2,000
43	cable	0,01446	0,133	6,039	2,000
44	camera	0,00281	0,044	6,083	0,667

45	carpet	0,00325	0,044	6,128	0,667
46	cassette	0,00620	0,044	6,172	0,667
47	CD(-ROM)	0,08768	1,066	7,238	16,000
48	CD(-ROM) player	0,00465	0,044	7,282	0,667
49	CD(-ROM) unit	0,00375	0,044	7,327	0,667
50	central processing unit	0,01287	0,089	7,416	1,333
51	charger	0,01187	0,133	7,549	2,000
52	chat	0,14459	1,465	9,014	22,000
53	chat room	0,00302	0,044	9,059	0,667
54	chip	0,00196	0,044	9,103	0,667
55	cinema	0,00403	0,044	9,147	0,667
56	Cinetube	0,00375	0,044	9,192	0,667
57	class	0,00137	0,044	9,236	0,667
58	click	0,04030	0,444	9,680	6,667
59	close	0,00375	0,044	9,725	0,667
60	code	0,00500	0,044	9,769	0,667
61	colour	0,00227	0,044	9,813	0,667
62	communication	0,01132	0,089	9,902	1,333
63	compact disc	0,02071	0,222	10,124	3,333
64	compile	0,00403	0,044	10,169	0,667
65	component	0,00433	0,044	10,213	0,667
66	computation	0,00325	0,044	10,258	0,667
67	computer	0,14403	1,110	11,368	16,667
68	computer game	0,00433	0,044	11,412	0,667
69	computer programmer	0,00196	0,044	11,456	0,667
70	computer science	0,00170	0,044	11,501	0,667
71	computer security	0,00403	0,044	11,545	0,667
72	computing task	0,00183	0,044	11,590	0,667
73	configuration	0,00262	0,044	11,634	0,667
74	connection	0,03194	0,400	12,034	6,000
75	contact	0,00158	0,044	12,078	0,667
76	control	0,00302	0,044	12,123	0,667
77	cookie	0,01253	0,133	12,256	2,000
78	copy	0,01214	0,222	12,478	3,333
79	CPU	0,02016	0,222	12,700	3,333
80	create	0,00196	0,044	12,744	0,667
81	cut	0,00183	0,044	12,789	0,667
82	data	0,00564	0,089	12,877	1,333
83	database	0,01215	0,178	13,055	2,667
84	delete	0,02053	0,222	13,277	3,333
85	design	0,01198	0,089	13,366	1,333
86	desk	0,01410	0,133	13,499	2,000
87	desktop	0,04615	0,400	13,899	6,000
88	desktop computer	0,00577	0,044	13,943	0,667
89	develop	0,00667	0,044	13,988	0,667
90	development	0,00577	0,044	14,032	0,667
91	dictionary	0,01052	0,133	14,165	2,000

92	directory	0,00302	0,044	14,210	0,667
93	disk	0,02170	0,311	14,520	4,667
94	diskette	0,00684	0,089	14,609	1,333
95	display	0,00577	0,044	14,654	0,667
96	document	0,01748	0,266	14,920	4,000
97	documentary	0,00302	0,044	14,964	0,667
98	domain	0,00170	0,044	15,009	0,667
99	dot	0,00433	0,044	15,053	0,667
100	download	0,16468	1,909	16,963	28,667
101	downloading	0,00537	0,044	17,007	0,667
102	drag	0,00158	0,044	17,052	0,667
103	drive	0,00433	0,044	17,096	0,667
104	Dropbox	0,00302	0,044	17,140	0,667
105	DVD	0,02691	0,444	17,584	6,667
106	DVD player	0,01618	0,178	17,762	2,667
107	dynamic random access memory	0,00281	0,044	17,806	0,667
108	eBay	0,00692	0,089	17,895	1,333
109	edit	0,00593	0,089	17,984	1,333
110	El Mundo	0,00403	0,044	18,028	0,667
111	El País	0,00736	0,089	18,117	1,333
112	electronic book	0,01077	0,089	18,206	1,333
113	electronic data processing	0,00281	0,044	18,250	0,667
114	electronic mail	0,00211	0,044	18,295	0,667
115	email	0,19997	2,176	20,471	32,667
116	eMule	0,01168	0,178	20,648	2,667
117	encryption	0,00170	0,044	20,693	0,667
118	enter	0,00678	0,089	20,782	1,333
119	entertainment	0,00183	0,044	20,826	0,667
120	erase	0,00508	0,089	20,915	1,333
121	error	0,00500	0,044	20,959	0,667
122	escape	0,00281	0,044	21,004	0,667
123	execute	0,01160	0,133	21,137	2,000
124	eXeLearning	0,00375	0,044	21,181	0,667
125	Facebook	0,27651	2,620	23,801	39,333
126	fake	0,00147	0,044	23,845	0,667
127	fax	0,00433	0,044	23,890	0,667
128	feed	0,00281	0,044	23,934	0,667
129	file	0,05978	0,844	24,778	12,667
130	fill	0,00302	0,044	24,822	0,667
131	film	0,02058	0,266	25,089	4,000
132	firewall	0,02839	0,355	25,444	5,333
133	fix	0,00349	0,044	25,488	0,667
134	flash memory	0,00577	0,044	25,533	0,667
135	floppy disk	0,04542	0,488	26,021	7,333
136	folder	0,04476	0,622	26,643	9,333
137	format	0,00885	0,133	26,776	2,000
138	forum	0,02459	0,311	27,087	4,667

139	free time	0,00325	0,044	27,131	0,667
140	friend	0,00302	0,044	27,176	0,667
141	front page	0,00244	0,044	27,220	0,667
142	function	0,00244	0,044	27,265	0,667
143	game	0,06338	0,666	27,931	10,000
144	gate	0,00325	0,044	27,975	0,667
145	gigabyte	0,00576	0,089	28,064	1,333
146	globalization	0,00620	0,044	28,108	0,667
147	Gmail	0,01963	0,222	28,330	3,333
148	Google	0,17235	1,687	30,018	25,333
149	Google Chrome	0,00281	0,044	30,062	0,667
150	Google Maps	0,00183	0,044	30,107	0,667
151	Google Reader	0,00620	0,044	30,151	0,667
152	Google Translate	0,00500	0,044	30,195	0,667
153	graphical display unit	0,01012	0,133	30,329	2,000
154	Grooveshark	0,00620	0,044	30,373	0,667
155	Groupon	0,00433	0,044	30,417	0,667
156	Guadalinex	0,00433	0,044	30,462	0,667
157	hacker	0,01025	0,178	30,639	2,667
158	hard disk	0,10958	1,288	31,927	19,333
159	hardware	0,19080	2,220	34,147	33,333
160	headphones	0,01008	0,133	34,281	2,000
161	help	0,00196	0,044	34,325	0,667
162	high definition	0,00375	0,044	34,369	0,667
163	holiday	0,00302	0,044	34,414	0,667
164	home	0,01529	0,133	34,547	2,000
165	Hot Potatoes	0,01664	0,178	34,725	2,667
166	Hotmail	0,06475	0,710	35,435	10,667
167	HP	0,00302	0,044	35,480	0,667
168	HTML	0,00500	0,044	35,524	0,667
169	hub	0,00627	0,089	35,613	1,333
170	huge	0,00375	0,044	35,657	0,667
171	hypertext	0,00349	0,044	35,702	0,667
172	icon	0,00196	0,044	35,746	0,667
173	ICT	0,00262	0,044	35,790	0,667
174	image	0,00782	0,089	35,879	1,333
175	InfoJobs	0,01121	0,089	35,968	1,333
176	informatics	0,00211	0,044	36,012	0,667
177	information	0,04306	0,533	36,545	8,000
178	information technology	0,00577	0,044	36,590	0,667
179	ink	0,00433	0,044	36,634	0,667
180	innovation	0,00302	0,044	36,679	0,667
181	input	0,00170	0,044	36,723	0,667
182	insert	0,00375	0,044	36,767	0,667
183	install	0,00736	0,089	36,856	1,333
184	installation	0,00281	0,044	36,901	0,667
185	instant messaging	0,00349	0,044	36,945	0,667

186	instruction	0,00537	0,044	36,989	0,667
187	interface	0,00325	0,044	37,034	0,667
188	Internet	0,11981	1,066	38,099	16,000
189	Internet Explorer	0,02490	0,266	38,366	4,000
190	Internet Protocol	0,00375	0,044	38,410	0,667
191	Interworking Function	0,00465	0,044	38,455	0,667
192	IP	0,00500	0,044	38,499	0,667
193	iPad	0,00325	0,044	38,544	0,667
194	ISP	0,00433	0,044	38,588	0,667
195	IT assistant	0,00403	0,044	38,632	0,667
196	Java	0,00965	0,089	38,721	1,333
197	JDownloader	0,00196	0,044	38,766	0,667
198	job	0,00211	0,044	38,810	0,667
199	joystick	0,00899	0,089	38,899	1,333
200	key	0,01138	0,178	39,076	2,667
201	keyboard	0,31772	2,753	41,829	41,333
202	keypad	0,00808	0,089	41,918	1,333
203	keyword	0,00529	0,089	42,007	1,333
204	knowledge	0,00147	0,044	42,052	0,667
205	label	0,00500	0,044	42,096	0,667
206	landscape format	0,00620	0,044	42,140	0,667
207	language	0,00244	0,044	42,185	0,667
208	laptop	0,21301	1,954	44,139	29,333
209	laser	0,00262	0,044	44,183	0,667
210	LCD screen	0,00227	0,044	44,227	0,667
211	LED	0,00227	0,044	44,272	0,667
212	letter	0,00211	0,044	44,316	0,667
213	link	0,04837	0,488	44,805	7,333
214	linkage	0,00302	0,044	44,849	0,667
215	Linux	0,02177	0,222	45,071	3,333
216	load	0,00355	0,089	45,160	1,333
217	lock	0,00880	0,089	45,249	1,333
218	log in	0,00603	0,089	45,337	1,333
219	log out	0,00128	0,044	45,382	0,667
220	login	0,00781	0,089	45,471	1,333
221	look for	0,00403	0,044	45,515	0,667
222	loudspeaker	0,00262	0,044	45,560	0,667
223	Macintosh	0,02509	0,266	45,826	4,000
224	mail	0,01333	0,089	45,915	1,333
225	mainframe computer	0,00433	0,044	45,959	0,667
226	malware	0,00375	0,044	46,004	0,667
227	mass media	0,00620	0,044	46,048	0,667
228	massively multiplayer online game	0,00211	0,044	46,092	0,667
229	maximize	0,00137	0,044	46,137	0,667
230	megabyte	0,00808	0,089	46,226	1,333

231	memory	0,01252	0,133	46,359	2,000
232	memory card	0,00732	0,133	46,492	2,000
233	Memory Stick	0,00795	0,133	46,625	2,000
234	message	0,00349	0,044	46,670	0,667
235	Messenger	0,05719	0,577	47,247	8,667
236	microphone	0,00281	0,044	47,291	0,667
237	Microsoft	0,02575	0,311	47,602	4,667
238	minimize	0,00147	0,044	47,647	0,667
239	modem	0,02628	0,355	48,002	5,333
240	monitor	0,06193	0,533	48,535	8,000
241	Moodle	0,00862	0,089	48,623	1,333
242	motherboard	0,02470	0,266	48,890	4,000
243	mouse	0,63362	5,373	54,263	80,667
244	mouse pad	0,00403	0,044	54,307	0,667
245	Mozilla Firefox	0,01525	0,178	54,485	2,667
246	MP3	0,00255	0,089	54,574	1,333
247	MS-DOS	0,00262	0,044	54,618	0,667
248	MSN	0,00537	0,044	54,663	0,667
249	multifunctional	0,00170	0,044	54,707	0,667
250	music	0,03248	0,355	55,062	5,333
251	My PC	0,00403	0,044	55,107	0,667
252	Myspace	0,00627	0,089	55,195	1,333
253	net	0,05973	0,533	55,728	8,000
254	netbook	0,02949	0,311	56,039	4,667
255	network	0,02490	0,266	56,306	4,000
256	network host	0,00500	0,044	56,350	0,667
257	networking	0,00667	0,044	56,394	0,667
258	new technologies	0,00281	0,044	56,439	0,667
259	news	0,01043	0,089	56,528	1,333
260	newsletter	0,00262	0,044	56,572	0,667
261	newspaper	0,00665	0,089	56,661	1,333
262	nickname	0,00244	0,044	56,705	0,667
263	notebook	0,00403	0,044	56,750	0,667
264	number	0,00379	0,089	56,838	1,333
265	object	0,00227	0,044	56,883	0,667
266	office suite	0,00375	0,044	56,927	0,667
267	online	0,02859	0,400	57,327	6,000
268	online banking	0,00302	0,044	57,371	0,667
269	online training	0,00244	0,044	57,416	0,667
270	open	0,00903	0,089	57,504	1,333
271	open source	0,00349	0,044	57,549	0,667
272	operating system	0,01968	0,266	57,815	4,000
273	option	0,00537	0,044	57,860	0,667
274	output	0,00158	0,044	57,904	0,667
275	page	0,02805	0,311	58,215	4,667
276	Paint	0,00929	0,089	58,304	1,333
277	parental control	0,00375	0,044	58,348	0,667

	software				
278	password	0,03995	0,488	58,837	7,333
279	paste	0,01174	0,222	59,059	3,333
280	PC	0,08129	0,710	59,769	10,667
281	PDF	0,00183	0,044	59,814	0,667
282	pendrive	0,12152	1,288	61,101	19,333
283	peripheral	0,00281	0,044	61,146	0,667
284	personal computer	0,04305	0,400	61,545	6,000
285	phone	0,00349	0,044	61,590	0,667
286	photo	0,01090	0,133	61,723	2,000
287	photography	0,00325	0,044	61,767	0,667
288	picture	0,00577	0,044	61,812	0,667
289	piracy	0,00537	0,044	61,856	0,667
290	pixel	0,00244	0,044	61,901	0,667
291	play	0,00886	0,089	61,989	1,333
292	plotter	0,00227	0,044	62,034	0,667
293	plug	0,00736	0,089	62,123	1,333
294	plug-in	0,00375	0,044	62,167	0,667
295	podcast	0,00677	0,089	62,256	1,333
296	pop-up menu	0,00667	0,044	62,300	0,667
297	porn page	0,00667	0,044	62,345	0,667
298	pornography	0,00781	0,089	62,433	1,333
299	portfolio	0,00375	0,044	62,478	0,667
300	post	0,00947	0,133	62,611	2,000
301	post a comment	0,00227	0,044	62,655	0,667
302	presentation	0,00244	0,044	62,700	0,667
303	press	0,00403	0,044	62,744	0,667
304	print	0,01179	0,133	62,877	2,000
305	printed circuit board	0,00349	0,044	62,922	0,667
306	printer	0,11325	1,243	64,165	18,667
307	processor	0,00128	0,044	64,210	0,667
308	profile	0,00433	0,044	64,254	0,667
309	program(me)	0,13715	1,776	66,030	26,667
310	programmer	0,00302	0,044	66,075	0,667
311	programming	0,00620	0,044	66,119	0,667
312	programming language	0,01000	0,089	66,208	1,333
313	protocol	0,00281	0,044	66,252	0,667
314	PSP	0,00620	0,044	66,297	0,667
315	RAM	0,02434	0,222	66,519	3,333
316	RAM disk	0,01583	0,178	66,696	2,667
317	reading	0,00440	0,089	66,785	1,333
318	reboot	0,00183	0,044	66,829	0,667
319	received	0,00302	0,044	66,874	0,667
320	record	0,00111	0,044	66,918	0,667
321	recorder	0,01464	0,222	67,140	3,333

322	recorderer	0,00375	0,044	67,185	0,667
323	rendering	0,00244	0,044	67,229	0,667
324	research	0,00662	0,089	67,318	1,333
325	reset	0,00349	0,044	67,362	0,667
326	resolution	0,00262	0,044	67,407	0,667
327	resource	0,00543	0,089	67,496	1,333
328	return	0,00325	0,044	67,540	0,667
329	review	0,00281	0,044	67,584	0,667
330	role play	0,00196	0,044	67,629	0,667
331	router	0,01780	0,222	67,851	3,333
332	run	0,01435	0,133	67,984	2,000
333	save	0,03018	0,400	68,384	6,000
334	scanner	0,06829	0,755	69,139	11,333
335	school	0,00500	0,044	69,183	0,667
336	screen	0,40740	3,819	73,002	57,333
337	screener	0,00403	0,044	73,046	0,667
338	screensaver	0,01250	0,133	73,179	2,000
339	search	0,04682	0,577	73,757	8,667
340	search engine	0,00183	0,044	73,801	0,667
341	search process	0,00325	0,044	73,845	0,667
342	security	0,00183	0,044	73,890	0,667
343	send	0,01309	0,222	74,112	3,333
344	serial	0,00244	0,044	74,156	0,667
345	series	0,00211	0,044	74,201	0,667
346	Series Yonkis	0,00903	0,089	74,290	1,333
347	server	0,01161	0,133	74,423	2,000
348	set up	0,00196	0,044	74,467	0,667
349	sex	0,00262	0,044	74,512	0,667
350	share	0,00196	0,044	74,556	0,667
351	share a file	0,00244	0,044	74,600	0,667
352	shift	0,00349	0,044	74,645	0,667
353	shopping	0,00648	0,089	74,734	1,333
354	signal	0,00433	0,044	74,778	0,667
355	site	0,04770	0,444	75,222	6,667
356	Skype	0,01685	0,222	75,444	3,333
357	slide	0,00500	0,044	75,488	0,667
358	SlideShare	0,00302	0,044	75,533	0,667
359	social media	0,01137	0,133	75,666	2,000
360	social network	0,09000	0,977	76,643	14,667
361	social networking website	0,00211	0,044	76,687	0,667
362	socket	0,00244	0,044	76,732	0,667
363	software	0,18819	2,220	78,952	33,333
364	solution	0,00390	0,089	79,041	1,333
365	sound	0,00211	0,044	79,085	0,667
366	sound card	0,00849	0,089	79,174	1,333

367	source	0,00374	0,089	79,263	1,333
368	spam	0,00651	0,089	79,352	1,333
369	speaker	0,01543	0,133	79,485	2,000
370	speed	0,00158	0,044	79,529	0,667
371	Spotify	0,00465	0,044	79,574	0,667
372	spreadsheet	0,00244	0,044	79,618	0,667
373	SPSS	0,00325	0,044	79,663	0,667
374	spyware	0,00403	0,044	79,707	0,667
375	start	0,01155	0,089	79,796	1,333
376	stop	0,00302	0,044	79,840	0,667
377	storage	0,00170	0,044	79,885	0,667
378	subscribe	0,00196	0,044	79,929	0,667
379	surf the Net	0,02832	0,266	80,195	4,000
380	switch off	0,00414	0,089	80,284	1,333
381	switch on	0,00420	0,089	80,373	1,333
382	system	0,01061	0,133	80,506	2,000
383	tab key	0,00325	0,044	80,551	0,667
384	table board	0,00500	0,044	80,595	0,667
385	tablet PC	0,01318	0,178	80,773	2,667
386	tag	0,00170	0,044	80,817	0,667
387	technology	0,00325	0,044	80,861	0,667
388	telephone line	0,00147	0,044	80,906	0,667
389	television	0,00281	0,044	80,950	0,667
390	terabyte	0,00536	0,089	81,039	1,333
391	text processor	0,00665	0,089	81,128	1,333
392	TFT-LCD screen	0,00211	0,044	81,172	0,667
393	The Sims	0,00403	0,044	81,217	0,667
394	The Times	0,00403	0,044	81,261	0,667
395	thread	0,00403	0,044	81,306	0,667
396	to google	0,00375	0,044	81,350	0,667
397	tool	0,00825	0,133	81,483	2,000
398	toolbar	0,00465	0,044	81,528	0,667
399	torrent file	0,00170	0,044	81,572	0,667
400	touchpad	0,00325	0,044	81,616	0,667
401	trackpad	0,00302	0,044	81,661	0,667
402	translator	0,00302	0,044	81,705	0,667
403	trash	0,00183	0,044	81,750	0,667
404	Trojan	0,01164	0,133	81,883	2,000
405	Tuenti	0,13667	1,421	83,304	21,333
406	tutorial	0,00227	0,044	83,348	0,667
407	Twitter	0,12870	1,199	84,547	18,000
408	typewriter	0,00325	0,044	84,591	0,667
409	typing	0,00667	0,044	84,636	0,667
410	under maintenance	0,00196	0,044	84,680	0,667
411	unit of energy	0,00465	0,044	84,725	0,667
412	Universal Serial Bus	0,00802	0,089	84,814	1,333

413	university	0,00170	0,044	84,858	0,667
414	unzip	0,00302	0,044	84,902	0,667
415	upload	0,06475	0,844	85,746	12,667
416	URL	0,00465	0,044	85,790	0,667
417	USB	0,04215	0,533	86,323	8,000
418	USB port	0,00211	0,044	86,368	0,667
419	useful	0,00196	0,044	86,412	0,667
420	user	0,00500	0,044	86,456	0,667
421	velocity	0,00302	0,044	86,501	0,667
422	video	0,04249	0,444	86,945	6,667
423	video call	0,00349	0,044	86,989	0,667
424	video card	0,00403	0,044	87,034	0,667
425	video game	0,02933	0,266	87,300	4,000
426	view	0,00302	0,044	87,345	0,667
427	virtual	0,00281	0,044	87,389	0,667
428	virus	0,05263	0,666	88,055	10,000
429	virus removal tool	0,00403	0,044	88,099	0,667
430	visual	0,00211	0,044	88,144	0,667
431	vocabulary	0,00147	0,044	88,188	0,667
432	wall	0,00196	0,044	88,233	0,667
433	web	0,21203	1,732	89,964	26,000
434	web application	0,00375	0,044	90,009	0,667
435	web browser	0,01334	0,133	90,142	2,000
436	web page	0,10590	1,110	91,252	16,667
437	webcam	0,05496	0,622	91,874	9,333
438	webcast	0,00262	0,044	91,918	0,667
439	WebQuest	0,00819	0,089	92,007	1,333
440	website	0,14396	1,377	93,384	20,667
441	widescreen	0,00349	0,044	93,428	0,667
442	Wi-Fi	0,09847	1,110	94,538	16,667
443	wiki	0,03149	0,266	94,805	4,000
444	WikiLeaks	0,01198	0,089	94,893	1,333
445	Wikipedia	0,07347	0,888	95,782	13,333
446	window	0,01121	0,089	95,870	1,333
447	Windows	0,11202	1,288	97,158	19,333
448	wire	0,01183	0,178	97,336	2,667
449	wireless	0,04031	0,488	97,824	7,333
450	wireless mouse	0,00325	0,044	97,869	0,667
451	wireless network	0,00183	0,044	97,913	0,667
452	WordPad	0,00147	0,044	97,957	0,667
453	Wordreference	0,00935	0,133	98,091	2,000
454	work	0,00678	0,089	98,179	1,333
455	World Wide Web	0,07234	0,755	98,934	11,333
456	worm	0,00302	0,044	98,979	0,667
457	write	0,00227	0,044	99,023	0,667
458	Yahoo	0,01642	0,222	99,245	3,333

459	Youtube	0,05943	0,710	99,956	10,667
460	zip	0,00325	0,044	100,000	0,667

2.2. Listado de palabras en inglés ordenado según el índice de disponibilidad

2.2.1. Human body parts

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	head	0,68183	4,862	4,862	83,333
2	eye	0,64489	5,056	9,918	86,667
3	leg	0,57612	4,979	14,897	85,333
4	arm	0,57588	4,706	19,603	80,667
5	nose	0,53836	4,823	24,426	82,667
6	finger	0,53468	5,056	29,483	86,667
7	hand	0,52240	4,590	34,072	78,667
8	mouth	0,50542	4,473	38,545	76,667
9	ear	0,40467	3,890	42,435	66,667
10	hair	0,38842	3,773	46,208	64,667
11	shoulder	0,25409	2,489	48,697	42,667
12	foot	0,25355	2,528	51,225	43,333
13	teeth	0,24580	2,723	53,948	46,667
14	neck	0,24475	2,606	56,554	44,667
15	nail	0,23941	2,800	59,354	48,000
16	feet	0,23043	2,178	61,532	37,333
17	toe	0,22448	2,606	64,138	44,667
18	knee	0,21975	2,684	66,822	46,000
19	lip	0,18584	2,100	68,923	36,000
20	heart	0,17242	2,178	71,101	37,333
21	face	0,15769	1,439	72,540	24,667
22	tongue	0,15497	1,945	74,485	33,333
23	back	0,13911	1,906	76,391	32,667
24	elbow	0,13626	1,322	77,713	22,667
25	eyebrow	0,11998	1,322	79,035	22,667
26	chest	0,08555	1,167	80,202	20,000
27	stomach	0,08103	1,245	81,447	21,333
28	wrist	0,07670	0,933	82,380	16,000
29	breast	0,07479	1,089	83,469	18,667
30	(eye)lash	0,07289	0,856	84,325	14,667
31	ankle	0,07161	0,817	85,142	14,000
32	skin	0,06846	0,817	85,959	14,000
33	cheek	0,05722	0,778	86,737	13,333
34	tooth	0,05488	0,622	87,359	10,667
35	ass	0,05433	0,700	88,059	12,000
36	bone	0,04860	0,739	88,798	12,667
37	brain	0,04849	0,545	89,343	9,333
38	thumb	0,04352	0,506	89,848	8,667
39	liver	0,03991	0,700	90,548	12,000
40	chin	0,03869	0,467	91,015	8,000
41	bottom	0,03806	0,622	91,637	10,667

42	belly	0,03617	0,428	92,065	7,333
43	kidney	0,03250	0,622	92,688	10,667
44	lung	0,03181	0,545	93,232	9,333
45	eyelid	0,03056	0,428	93,660	7,333
46	hip	0,03051	0,350	94,010	6,000
47	penis	0,02487	0,350	94,360	6,000
48	muscle	0,02318	0,467	94,827	8,000
49	forehead	0,01748	0,272	95,099	4,667
50	body	0,01697	0,194	95,294	3,333
51	forearm	0,01640	0,194	95,488	3,333
52	blood	0,01534	0,194	95,683	3,333
53	jaw	0,01394	0,194	95,877	3,333
54	vagina	0,01262	0,233	96,110	4,000
55	throat	0,01207	0,194	96,305	3,333
56	vein	0,01147	0,233	96,538	4,000
57	belly button	0,01138	0,156	96,694	2,667
58	spleen	0,01137	0,117	96,811	2,000
59	waist	0,01077	0,194	97,005	3,333
60	skull	0,00990	0,117	97,122	2,000
61	mind	0,00912	0,078	97,200	1,333
62	nostril	0,00880	0,078	97,277	1,333
63	wrinkle	0,00804	0,078	97,355	1,333
64	ear lobe	0,00735	0,117	97,472	2,000
65	physical	0,00667	0,039	97,511	0,667
66	calf muscle	0,00644	0,117	97,627	2,000
67	pussy	0,00631	0,078	97,705	1,333
68	bust	0,00602	0,078	97,783	1,333
69	tummy	0,00593	0,078	97,861	1,333
70	food	0,00577	0,039	97,900	0,667
71	dick	0,00576	0,078	97,977	1,333
72	palm	0,00537	0,039	98,016	0,667
73	butt	0,00525	0,078	98,094	1,333
74	armpit	0,00508	0,078	98,172	1,333
75	cheer	0,00500	0,039	98,211	0,667
76	limb	0,00500	0,039	98,250	0,667
77	trunk	0,00500	0,039	98,289	0,667
78	heel	0,00465	0,039	98,327	0,667
79	biceps	0,00450	0,078	98,405	1,333
80	pelvis	0,00433	0,039	98,444	0,667
81	intestine	0,00410	0,078	98,522	1,333
82	larynx	0,00410	0,078	98,600	1,333
83	brow	0,00385	0,078	98,678	1,333
84	epidermis	0,00349	0,039	98,716	0,667
85	bob	0,00325	0,039	98,755	0,667
86	cranium	0,00325	0,039	98,794	0,667
87	fist	0,00325	0,039	98,833	0,667
88	joint	0,00325	0,039	98,872	0,667
89	pupil	0,00325	0,039	98,911	0,667
90	bum	0,00302	0,039	98,950	0,667

91	bump	0,00302	0,039	98,989	0,667
92	cunt	0,00302	0,039	99,028	0,667
93	iris	0,00302	0,039	99,067	0,667
94	tip	0,00302	0,039	99,105	0,667
95	abdomen	0,00281	0,039	99,144	0,667
96	triceps	0,00281	0,039	99,183	0,667
97	artery	0,00262	0,039	99,222	0,667
98	Stone	0,00248	0,078	99,300	1,333
99	hear	0,00229	0,078	99,378	1,333
100	chick	0,00227	0,039	99,417	0,667
101	cock	0,00227	0,039	99,455	0,667
102	axis	0,00183	0,039	99,494	0,667
103	ball	0,00183	0,039	99,533	0,667
104	groin	0,00183	0,039	99,572	0,667
105	elbow joint	0,00147	0,039	99,611	0,667
106	hamstring	0,00147	0,039	99,650	0,667
107	column	0,00137	0,039	99,689	0,667
108	knuckle	0,00137	0,039	99,728	0,667
109	testicle	0,00137	0,039	99,767	0,667
110	thigh	0,00137	0,039	99,806	0,667
111	illness	0,00128	0,039	99,844	0,667
112	blood vessel	0,00111	0,039	99,883	0,667
113	tit	0,00111	0,039	99,922	0,667
114	navel	0,00077	0,039	99,961	0,667
115	nipple	0,00072	0,039	100,000	0,667

2.2.2. Clothes

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	trouser	0,71442	6,3523	6,3523	88,0000
2	T-shirt	0,66029	5,9673	12,3195	82,6667
3	skirt	0,55935	5,2454	17,5650	72,6667
4	shoe	0,53759	5,7748	23,3397	80,0000
5	jeans	0,50265	4,9086	28,2483	68,0000
6	shirt	0,46558	4,3792	32,6275	60,6667
7	sock	0,40157	4,7642	37,3917	66,0000
8	jacket	0,34890	3,8017	41,1935	52,6667
9	hat	0,30655	3,4649	44,6583	48,0000
10	dress	0,27882	3,2724	47,9307	45,3333
11	scarf	0,23447	2,7430	50,6737	38,0000
12	coat	0,19665	2,4543	53,1280	34,0000
13	pullover	0,19566	2,1655	55,2936	30,0000
14	boot	0,17799	2,4543	57,7478	34,0000
15	glove	0,17624	2,1655	59,9134	30,0000
16	shorts	0,17067	2,0212	61,9346	28,0000
17	bra	0,15441	2,2137	64,1482	30,6667
18	sweater	0,14538	1,6362	65,7844	22,6667
19	tie	0,13988	1,9249	67,7093	26,6667
20	underwear	0,13586	1,7806	69,4899	24,6667
21	pants	0,13574	1,4918	70,9817	20,6667
22	jumper	0,13067	1,2031	72,1848	16,6667
23	suit	0,12693	1,6843	73,8691	23,3333
24	blouse	0,11204	1,2993	75,1684	18,0000
25	jersey	0,09469	1,1550	76,3234	16,0000
26	belt	0,08934	1,3474	77,6708	18,6667
27	panties	0,08272	1,1068	78,7777	15,3333
28	cardigan	0,08047	0,9143	79,6920	12,6667
29	trainer	0,07278	0,9143	80,6064	12,6667
30	cap	0,06774	0,9625	81,5688	13,3333
31	leggings	0,05682	0,7700	82,3388	10,6667
32	slip	0,05222	0,7218	83,0606	10,0000
33	tracksuit	0,04950	0,7218	83,7825	10,0000
34	anorak	0,04851	0,6256	84,4081	8,6667
35	knickers	0,04811	0,6256	85,0337	8,6667
36	pyjamas	0,04720	0,7700	85,8037	10,6667
37	raincoat	0,04274	0,7218	86,5255	10,0000
38	boxer shorts	0,03966	0,4331	86,9586	6,0000
39	top	0,03200	0,4331	87,3917	6,0000
40	cupboard	0,03002	0,3850	87,7767	5,3333
41	glasses	0,02994	0,3369	88,1136	4,6667
42	high heels	0,02852	0,4331	88,5467	6,0000
43	bag	0,02748	0,2887	88,8354	4,0000
44	tights	0,02648	0,4331	89,2685	6,0000
45	sandal	0,02175	0,2887	89,5573	4,0000

46	foulard	0,02118	0,2887	89,8460	4,0000
47	flip-flop	0,01903	0,3369	90,1829	4,6667
48	button	0,01866	0,2406	90,4235	3,3333
49	heel	0,01859	0,3850	90,8085	5,3333
50	slipper	0,01832	0,3369	91,1453	4,6667
51	vest	0,01746	0,2406	91,3859	3,3333
52	smoking jacket	0,01666	0,2406	91,6266	3,3333
53	umbrella	0,01655	0,2406	91,8672	3,3333
54	miniskirt	0,01613	0,2406	92,1078	3,3333
55	thong	0,01431	0,2406	92,3484	3,3333
56	necklace	0,01430	0,2887	92,6372	4,0000
57	stiletto	0,01418	0,1925	92,8296	2,6667
58	stocking	0,01404	0,1444	92,9740	2,0000
59	sneaker	0,01330	0,2406	93,2146	3,3333
60	sunglasses	0,01325	0,1444	93,3590	2,0000
61	mini-short	0,01293	0,1925	93,5515	2,6667
62	bikini	0,01221	0,1925	93,7440	2,6667
63	baseball cap	0,01164	0,1444	93,8884	2,0000
64	cotton	0,01143	0,1444	94,0327	2,0000
65	blazer	0,01103	0,1444	94,1771	2,0000
66	sweatshirt	0,01042	0,0962	94,2733	1,3333
67	apron	0,01032	0,1444	94,4177	2,0000
68	trunks	0,01023	0,1444	94,5621	2,0000
69	tennis shoe	0,00862	0,1444	94,7064	2,0000
70	shop	0,00817	0,0962	94,8027	1,3333
71	underpants	0,00808	0,0962	94,8989	1,3333
72	ring	0,00807	0,1444	95,0433	2,0000
73	sports jacket	0,00806	0,0962	95,1396	1,3333
74	high-heeled shoe	0,00802	0,0962	95,2358	1,3333
75	chinos	0,00758	0,0962	95,3320	1,3333
76	swimsuit	0,00756	0,1444	95,4764	2,0000
77	robe	0,00752	0,0962	95,5727	1,3333
78	overall	0,00740	0,1444	95,7170	2,0000
79	watch	0,00700	0,0962	95,8133	1,3333
80	Hood	0,00665	0,0962	95,9095	1,3333
81	lingerie	0,00647	0,0962	96,0058	1,3333
82	earring	0,00640	0,1444	96,1501	2,0000
83	hose	0,00620	0,0481	96,1983	0,6667
84	nightgown	0,00605	0,0962	96,2945	1,3333
85	handkerchief	0,00586	0,0962	96,3908	1,3333
86	shopping centre	0,00577	0,0481	96,4389	0,6667
87	turtleneck	0,00564	0,0962	96,5351	1,3333
88	bracelet	0,00540	0,0962	96,6314	1,3333
89	chemise	0,00537	0,0481	96,6795	0,6667
90	knit	0,00537	0,0481	96,7276	0,6667
91	shoulder pad	0,00537	0,0481	96,7757	0,6667
92	necktie	0,00500	0,0481	96,8239	0,6667
93	wool	0,00500	0,0481	96,8720	0,6667
94	zip	0,00500	0,0481	96,9201	0,6667

95	sleeve	0,00496	0,0962	97,0164	1,3333
96	athletic shoe	0,00471	0,0962	97,1126	1,3333
97	accessories	0,00465	0,0481	97,1607	0,6667
98	body stocking	0,00465	0,0481	97,2089	0,6667
99	smart	0,00465	0,0481	97,2570	0,6667
100	swimming trunks	0,00465	0,0481	97,3051	0,6667
101	jumpsuit	0,00432	0,0962	97,4013	1,3333
102	bathrobe	0,00414	0,0962	97,4976	1,3333
103	swimming costume	0,00409	0,0962	97,5938	1,3333
104	wig	0,00403	0,0481	97,6420	0,6667
105	bathing suit	0,00375	0,0481	97,6901	0,6667
106	easy wear	0,00375	0,0481	97,7382	0,6667
107	foot	0,00375	0,0481	97,7863	0,6667
108	handbag	0,00375	0,0481	97,8345	0,6667
109	lace	0,00375	0,0481	97,8826	0,6667
110	pencil skirt	0,00375	0,0481	97,9307	0,6667
111	short	0,00375	0,0481	97,9788	0,6667
112	tuxedo	0,00364	0,0962	98,0751	1,3333
113	bell-bottoms	0,00349	0,0481	98,1232	0,6667
114	casual wear	0,00349	0,0481	98,1713	0,6667
115	purse	0,00349	0,0481	98,2194	0,6667
116	bomber jacket	0,00325	0,0481	98,2676	0,6667
117	bow tie	0,00325	0,0481	98,3157	0,6667
118	headscarf	0,00325	0,0481	98,3638	0,6667
119	mitten	0,00322	0,0962	98,4601	1,3333
120	French sneaker	0,00302	0,0481	98,5082	0,6667
121	cloak	0,00281	0,0481	98,5563	0,6667
122	hotpants	0,00281	0,0481	98,6044	0,6667
123	nightdress	0,00281	0,0481	98,6526	0,6667
124	waist	0,00281	0,0481	98,7007	0,6667
125	waistcoat	0,00281	0,0481	98,7488	0,6667
126	money	0,00262	0,0481	98,7969	0,6667
127	sleepers	0,00262	0,0481	98,8450	0,6667
128	dressing gown	0,00244	0,0481	98,8932	0,6667
129	collar	0,00227	0,0481	98,9413	0,6667
130	mackintosh	0,00227	0,0481	98,9894	0,6667
131	cocktail dress	0,00211	0,0481	99,0375	0,6667
132	gaberline	0,00211	0,0481	99,0857	0,6667
133	gown	0,00211	0,0481	99,1338	0,6667
134	blanket	0,00196	0,0481	99,1819	0,6667
135	wool shirt	0,00196	0,0481	99,2300	0,6667
136	body	0,00183	0,0481	99,2782	0,6667
137	flat sandal	0,00183	0,0481	99,3263	0,6667
138	handcuffs	0,00183	0,0481	99,3744	0,6667
139	bonnet	0,00170	0,0481	99,4225	0,6667
140	rompers	0,00170	0,0481	99,4706	0,6667
141	briefs	0,00158	0,0481	99,5188	0,6667
142	designer	0,00158	0,0481	99,5669	0,6667
143	dungarees	0,00158	0,0481	99,6150	0,6667

144	swimming cap	0,00158	0,0481	99,6631	0,6667
145	style	0,00147	0,0481	99,7113	0,6667
146	undergarment	0,00147	0,0481	99,7594	0,6667
147	parka	0,00137	0,0481	99,8075	0,6667
148	towel	0,00137	0,0481	99,8556	0,6667
149	fleece jacket	0,00111	0,0481	99,9038	0,6667
150	cuff	0,00103	0,0481	99,9519	0,6667
151	clog	0,00096	0,0481	100,0000	0,6667

2.2.3. Food and drink

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	wáter	0,56424	3,884	3,884	75,333
2	meat	0,40602	2,888	6,772	56,000
3	fish	0,39906	3,128	9,900	60,667
4	tomato	0,35255	3,163	13,063	61,333
5	potato	0,32813	2,750	15,813	53,333
6	wine	0,30385	2,303	18,116	44,667
7	milk	0,29564	2,372	20,488	46,000
8	Coke	0,29469	2,200	22,688	42,667
9	orange	0,28796	2,338	25,026	45,333
10	juice	0,25189	2,166	27,191	42,000
11	beer	0,24799	1,891	29,082	36,667
12	bread	0,24188	1,994	31,076	38,667
13	chip	0,22958	1,891	32,967	36,667
14	coffee	0,22335	2,097	35,064	40,667
15	Apple	0,20598	1,788	36,851	34,667
16	vegetable	0,19399	1,409	38,261	27,333
17	hamburger	0,18764	1,409	39,670	27,333
18	banana	0,18259	1,650	41,320	32,000
19	chicken	0,16837	1,581	42,901	30,667
20	pizza	0,16227	1,272	44,173	24,667
21	tea	0,15251	1,444	45,617	28,000
22	rice	0,15020	1,169	46,786	22,667
23	egg	0,14981	1,169	47,955	22,667
24	salad	0,14705	1,375	49,330	26,667
25	lettuce	0,14051	1,375	50,705	26,667
26	soup	0,13753	1,169	51,873	22,667
27	onion	0,13747	1,238	53,111	24,000
28	carrot	0,13388	1,238	54,349	24,000
29	lemon	0,13140	1,203	55,552	23,333
30	cake	0,11701	1,100	56,652	21,333
31	fruit	0,11622	0,928	57,580	18,000
32	sugar	0,10577	1,134	58,714	22,000
33	salt	0,10292	1,169	59,883	22,667
34	sándwich	0,10083	0,859	60,743	16,667
35	strawberry	0,09864	1,031	61,774	20,000
36	cheese	0,09440	0,963	62,736	18,667
37	bean	0,09278	0,859	63,596	16,667
38	pepper	0,09183	0,997	64,593	19,333
39	chocolate	0,09008	0,928	65,521	18,000
40	whisky	0,08990	0,825	66,346	16,000
41	beef	0,08628	0,791	67,136	15,333
42	watermelon	0,08564	0,859	67,996	16,667
43	pasta	0,08330	0,722	68,718	14,000
44	spaghetti	0,08310	0,584	69,302	11,333
45	pineapple	0,08196	0,756	70,058	14,667

46	oil	0,08142	0,963	71,021	18,667
47	butter	0,07932	0,756	71,777	14,667
48	steak	0,07124	0,619	72,396	12,000
49	sausage	0,06795	0,722	73,118	14,000
50	orange juice	0,06745	0,584	73,702	11,333
51	pear	0,06695	0,688	74,390	13,333
52	pork	0,06601	0,688	75,077	13,333
53	soda	0,05878	0,447	75,524	8,667
54	jam	0,05529	0,619	76,143	12,000
55	Coca-Cola	0,05450	0,309	76,452	6,000
56	tuna	0,05059	0,619	77,071	12,000
57	sauce	0,04783	0,447	77,518	8,667
58	omelette	0,04776	0,378	77,896	7,333
59	melón	0,04754	0,584	78,481	11,333
60	lamb	0,04606	0,481	78,962	9,333
61	ice cream	0,04434	0,550	79,512	10,667
62	lemonade	0,04331	0,447	79,959	8,667
63	yogurt	0,03903	0,447	80,406	8,667
64	cucumber	0,03657	0,378	80,784	7,333
65	rum	0,03460	0,344	81,128	6,667
66	garlic	0,03295	0,413	81,540	8,000
67	grape	0,03192	0,344	81,884	6,667
68	ham	0,03187	0,378	82,262	7,333
69	alcohol	0,03136	0,241	82,503	4,667
70	Burger	0,03119	0,275	82,778	5,333
71	cod	0,03105	0,309	83,087	6,000
72	pea	0,03022	0,378	83,465	7,333
73	biscuit	0,03009	0,275	83,740	5,333
74	meal	0,02917	0,172	83,912	3,333
75	pie	0,02888	0,309	84,221	6,000
76	cookie	0,02821	0,344	84,565	6,667
77	meatball	0,02811	0,241	84,806	4,667
78	milkshake	0,02776	0,275	85,081	5,333
79	spirit	0,02774	0,241	85,321	4,667
80	paella	0,02583	0,172	85,493	3,333
81	soft drink	0,02393	0,206	85,700	4,000
82	peach	0,02388	0,309	86,009	6,000
83	seafood	0,02215	0,206	86,215	4,000
84	liquor	0,02208	0,206	86,421	4,000
85	chicken nugget	0,02146	0,138	86,559	2,667
86	vinegar	0,02119	0,275	86,834	5,333
87	gin	0,02005	0,172	87,006	3,333
88	sweet	0,01985	0,275	87,281	5,333
89	snack	0,01788	0,138	87,418	2,667
90	cereal	0,01746	0,206	87,625	4,000
91	bacon	0,01739	0,206	87,831	4,000
92	olive	0,01739	0,206	88,037	4,000
93	aubergine	0,01738	0,206	88,243	4,000
94	stew	0,01710	0,172	88,415	3,333

95	beefsteak	0,01675	0,103	88,518	2,000
96	vodka	0,01633	0,172	88,690	3,333
97	cherry	0,01548	0,172	88,862	3,333
98	cauliflower	0,01541	0,172	89,034	3,333
99	ketchup	0,01518	0,138	89,172	2,667
100	cider	0,01477	0,138	89,309	2,667
101	potato crisp	0,01400	0,103	89,412	2,000
102	marmalade	0,01387	0,138	89,550	2,667
103	prawn	0,01363	0,172	89,722	3,333
104	salmon	0,01340	0,172	89,893	3,333
105	avocado	0,01334	0,172	90,065	3,333
106	dessert	0,01312	0,138	90,203	2,667
107	red wine	0,01277	0,138	90,340	2,667
108	white wine	0,01218	0,103	90,443	2,000
109	toast	0,01209	0,138	90,581	2,667
110	popcorn	0,01202	0,138	90,718	2,667
111	corn	0,01196	0,172	90,890	3,333
112	honey	0,01190	0,103	90,993	2,000
113	lentil	0,01168	0,103	91,097	2,000
114	chop	0,01152	0,103	91,200	2,000
115	cornflakes	0,01144	0,103	91,303	2,000
116	hot dog	0,01137	0,138	91,440	2,667
117	cabbage	0,01115	0,069	91,509	1,333
118	mayonnaise	0,01101	0,138	91,647	2,667
119	broccoli	0,01101	0,138	91,784	2,667
120	roast	0,01097	0,103	91,887	2,000
121	mushroom	0,01082	0,138	92,025	2,667
122	take away	0,01077	0,069	92,094	1,333
123	drink	0,01072	0,069	92,162	1,333
124	olive oil	0,01034	0,138	92,300	2,667
125	octopus	0,01008	0,138	92,437	2,667
126	fast food	0,01007	0,069	92,506	1,333
127	peanut	0,01001	0,138	92,644	2,667
128	Fanta	0,00997	0,103	92,747	2,000
129	kiwi	0,00969	0,103	92,850	2,000
130	nut	0,00963	0,069	92,919	1,333
131	pork chop	0,00959	0,069	92,987	1,333
132	roast chicken	0,00934	0,069	93,056	1,333
133	cupcake	0,00925	0,069	93,125	1,333
134	eggplant	0,00899	0,103	93,228	2,000
135	fish and chips	0,00891	0,103	93,331	2,000
136	ice	0,00873	0,069	93,400	1,333
137	Pepsi	0,00873	0,069	93,469	1,333
138	duck	0,00835	0,069	93,537	1,333
139	flour	0,00816	0,069	93,606	1,333
140	mojito	0,00806	0,069	93,675	1,333
141	pumpkin	0,00805	0,069	93,744	1,333
142	coconut	0,00799	0,069	93,812	1,333
143	alcopop	0,00788	0,103	93,915	2,000

144	rabbit	0,00786	0,069	93,984	1,333
145	plate	0,00779	0,103	94,087	2,000
146	crisp	0,00763	0,069	94,156	1,333
147	restaurant	0,00757	0,103	94,259	2,000
148	sole	0,00751	0,069	94,328	1,333
149	fried	0,00738	0,069	94,397	1,333
150	cocktail	0,00725	0,069	94,465	1,333
151	squid	0,00714	0,103	94,569	2,000
152	Berry	0,00694	0,069	94,637	1,333
153	pig	0,00694	0,069	94,706	1,333
154	noodle	0,00685	0,069	94,775	1,333
155	chili	0,00667	0,034	94,809	0,667
156	gin and tonic	0,00667	0,034	94,844	0,667
157	Indian food	0,00667	0,034	94,878	0,667
158	McDonald's	0,00667	0,034	94,912	0,667
159	sushi	0,00653	0,069	94,981	1,333
160	snail	0,00649	0,069	95,050	1,333
161	French fries	0,00628	0,069	95,119	1,333
162	monkfish	0,00626	0,034	95,153	0,667
163	penne	0,00626	0,034	95,187	0,667
164	refresher	0,00626	0,034	95,222	0,667
165	Diet Coke	0,00613	0,069	95,290	1,333
166	veal	0,00613	0,069	95,359	1,333
167	cranberry	0,00589	0,103	95,462	2,000
168	tangerine	0,00589	0,034	95,497	0,667
169	Candy	0,00582	0,103	95,600	2,000
170	sparkling water	0,00562	0,069	95,669	1,333
171	alcoholic drink	0,00553	0,034	95,703	0,667
172	cola	0,00553	0,034	95,737	0,667
173	mineral water	0,00553	0,034	95,772	0,667
174	oxtail soup	0,00553	0,034	95,806	0,667
175	shellfish	0,00553	0,034	95,840	0,667
176	traditional food	0,00553	0,034	95,875	0,667
177	glass	0,00547	0,103	95,978	2,000
178	roast beef	0,00531	0,069	96,047	1,333
179	foie gras	0,00520	0,034	96,081	0,667
180	gazpacho	0,00520	0,034	96,116	0,667
181	joint	0,00520	0,034	96,150	0,667
182	raspberry	0,00520	0,034	96,184	0,667
183	codfish	0,00515	0,069	96,253	1,333
184	fruit juice	0,00489	0,034	96,287	0,667
185	ginger	0,00489	0,034	96,322	0,667
186	jelly	0,00489	0,034	96,356	0,667
187	nacho	0,00489	0,034	96,391	0,667
188	pudding	0,00489	0,034	96,425	0,667
189	turnip	0,00489	0,034	96,459	0,667
190	blueberry	0,00475	0,069	96,528	1,333
191	crab	0,00459	0,034	96,562	0,667
192	hake	0,00459	0,034	96,597	0,667

193	anchovy	0,00431	0,034	96,631	0,667
194	champagne	0,00431	0,034	96,666	0,667
195	coleslaw	0,00431	0,034	96,700	0,667
196	fresh milk	0,00431	0,034	96,734	0,667
197	grapefruit	0,00431	0,034	96,769	0,667
198	oat	0,00431	0,034	96,803	0,667
199	sparkling wine	0,00431	0,034	96,837	0,667
200	cream	0,00426	0,069	96,906	1,333
201	BBQ	0,00405	0,034	96,941	0,667
202	cow	0,00405	0,034	96,975	0,667
203	hot chocolate	0,00405	0,034	97,009	0,667
204	sardine	0,00405	0,034	97,044	0,667
205	shrimp	0,00405	0,034	97,078	0,667
206	Yorkshire pudding	0,00405	0,034	97,112	0,667
207	cinnamon	0,00402	0,069	97,181	1,333
208	baked potato	0,00381	0,034	97,216	0,667
209	grab	0,00381	0,034	97,250	0,667
210	gravy	0,00381	0,034	97,284	0,667
211	well-cooked	0,00381	0,034	97,319	0,667
212	barley	0,00358	0,034	97,353	0,667
213	chamomile tea	0,00358	0,034	97,387	0,667
214	energy drink	0,00358	0,034	97,422	0,667
215	grapefruit juice	0,00358	0,034	97,456	0,667
216	mussel	0,00358	0,034	97,491	0,667
217	paste	0,00358	0,034	97,525	0,667
218	rare	0,00358	0,034	97,559	0,667
219	spice	0,00358	0,034	97,594	0,667
220	spoon	0,00358	0,034	97,628	0,667
221	Chinese food	0,00356	0,069	97,697	1,333
222	mandarin	0,00337	0,069	97,766	1,333
223	fork	0,00336	0,034	97,800	0,667
224	legume	0,00336	0,034	97,834	0,667
225	rib eye	0,00336	0,034	97,869	0,667
226	turkey	0,00336	0,034	97,903	0,667
227	light	0,00316	0,034	97,937	0,667
228	plum pudding	0,00316	0,034	97,972	0,667
229	rib	0,00316	0,034	98,006	0,667
230	tortellini	0,00316	0,034	98,041	0,667
231	almond	0,00297	0,034	98,075	0,667
232	cordial	0,00297	0,034	98,109	0,667
233	cup	0,00297	0,034	98,144	0,667
234	fig	0,00297	0,034	98,178	0,667
235	poultry	0,00297	0,034	98,212	0,667
236	ravioli	0,00297	0,034	98,247	0,667
237	wasabi	0,00297	0,034	98,281	0,667
238	cheesecake	0,00279	0,034	98,316	0,667
239	spring wáter	0,00279	0,034	98,350	0,667
240	bar	0,00262	0,034	98,384	0,667
241	walnut	0,00262	0,034	98,419	0,667

242	black tea	0,00247	0,034	98,453	0,667
243	brandy	0,00247	0,034	98,487	0,667
244	chestnut	0,00247	0,034	98,522	0,667
245	salami	0,00247	0,034	98,556	0,667
246	smoothie	0,00247	0,034	98,591	0,667
247	clam	0,00232	0,034	98,625	0,667
248	ginger ale	0,00232	0,034	98,659	0,667
249	pancake	0,00232	0,034	98,694	0,667
250	red tea	0,00232	0,034	98,728	0,667
251	gooseberry	0,00218	0,034	98,762	0,667
252	Kir	0,00218	0,034	98,797	0,667
253	table	0,00218	0,034	98,831	0,667
254	tomato sauce	0,00218	0,034	98,866	0,667
255	white tea	0,00218	0,034	98,900	0,667
256	blackberry	0,00205	0,034	98,934	0,667
257	cottage pie	0,00205	0,034	98,969	0,667
258	green tea	0,00205	0,034	99,003	0,667
259	Kir Royal	0,00205	0,034	99,037	0,667
260	muffin	0,00205	0,034	99,072	0,667
261	sangría	0,00205	0,034	99,106	0,667
262	veggie	0,00205	0,034	99,141	0,667
263	cappuccino	0,00192	0,034	99,175	0,667
264	paprika	0,00192	0,034	99,209	0,667
265	cognac	0,00181	0,034	99,244	0,667
266	still wáter	0,00181	0,034	99,278	0,667
267	crème caramel	0,00170	0,034	99,312	0,667
268	plum	0,00170	0,034	99,347	0,667
269	sunflower seed oil	0,00170	0,034	99,381	0,667
270	bagel	0,00160	0,034	99,416	0,667
271	bread roll	0,00160	0,034	99,450	0,667
272	caramel	0,00160	0,034	99,484	0,667
273	panna cotta	0,00160	0,034	99,519	0,667
274	scampi	0,00160	0,034	99,553	0,667
275	comfort food	0,00150	0,034	99,587	0,667
276	starter	0,00150	0,034	99,622	0,667
277	crepe	0,00141	0,034	99,656	0,667
278	main course	0,00141	0,034	99,691	0,667
279	leek	0,00132	0,034	99,725	0,667
280	artichoke	0,00124	0,034	99,759	0,667
281	blackcurrant	0,00124	0,034	99,794	0,667
282	hum	0,00124	0,034	99,828	0,667
283	meat loaf	0,00117	0,034	99,862	0,667
284	brisket	0,00110	0,034	99,897	0,667
285	weir	0,00091	0,034	99,931	0,667
286	mustard	0,00086	0,034	99,966	0,667
287	bottle	0,00081	0,034	100,000	0,667

2.2.4. The school (furniture and school material)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	pencil	0,75392	6,965	6,965	98,667
2	table	0,66936	5,741	12,706	81,333
3	chair	0,64510	5,788	18,494	82,000
4	pen	0,54858	5,224	23,718	74,000
5	book	0,48006	5,271	28,988	74,667
6	blackboard	0,47853	4,376	33,365	62,000
7	desk	0,31190	2,824	36,188	40,000
8	computer	0,30757	3,953	40,141	56,000
9	notebook	0,30747	3,624	43,765	51,333
10	rubber	0,29226	3,294	47,059	46,667
11	teacher	0,25663	2,682	49,741	38,000
12	paper	0,20505	2,447	52,188	34,667
13	window	0,18051	2,259	54,447	32,000
14	chalk	0,14936	1,835	56,282	26,000
15	ruler	0,14245	1,553	57,835	22,000
16	door	0,13374	1,694	59,529	24,000
17	rule	0,12303	1,365	60,894	19,333
18	pupil	0,10986	1,224	62,118	17,333
19	student	0,10003	1,176	63,294	16,667
20	eraser	0,09972	1,318	64,612	18,667
21	classroom	0,09871	1,129	65,741	16,000
22	bag	0,09744	1,271	67,012	18,000
23	pencil case	0,08178	1,082	68,094	15,333
24	map	0,07756	1,082	69,176	15,333
25	workbook	0,07543	0,894	70,071	12,667
26	board	0,06466	0,706	70,776	10,000
27	dictionary	0,05989	0,847	71,624	12,000
28	ballpoint pen	0,05836	0,612	72,235	8,667
29	wardrobe	0,05209	0,659	72,894	9,333
30	gym	0,05160	0,706	73,600	10,000
31	sheet	0,05139	0,612	74,212	8,667
32	folder	0,04611	0,612	74,824	8,667
33	projector	0,04271	0,518	75,341	7,333
34	exam	0,04219	0,565	75,906	8,000
35	schoolbag	0,03998	0,612	76,518	8,667
36	class	0,03795	0,424	76,941	6,000
37	sharpener	0,03658	0,471	77,412	6,667
38	bin	0,03645	0,471	77,882	6,667
39	color	0,03395	0,424	78,306	6,000
40	television	0,03150	0,424	78,729	6,000
41	timetable	0,03071	0,376	79,106	5,333
42	subject	0,02924	0,424	79,529	6,000
43	whiteboard	0,02916	0,282	79,812	4,000
44	professor	0,02844	0,282	80,094	4,000
45	playground	0,02688	0,329	80,424	4,667

46	maths	0,02682	0,282	80,706	4,000
47	carpet	0,02581	0,329	81,035	4,667
48	screen	0,02532	0,329	81,365	4,667
49	hanger	0,02516	0,329	81,694	4,667
50	library	0,02404	0,329	82,024	4,667
51	student book	0,02249	0,329	82,353	4,667
52	toilet	0,02184	0,282	82,635	4,000
53	marker	0,02105	0,329	82,965	4,667
54	notepad	0,02028	0,282	83,247	4,000
55	child	0,02018	0,188	83,435	2,667
56	page	0,01896	0,188	83,624	2,667
57	break	0,01885	0,282	83,906	4,000
58	lamp	0,01880	0,329	84,235	4,667
59	note	0,01873	0,235	84,471	3,333
60	cupboard	0,01871	0,235	84,706	3,333
61	poster	0,01869	0,235	84,941	3,333
62	mouse	0,01868	0,329	85,271	4,667
63	textbook	0,01812	0,188	85,459	2,667
64	calculator	0,01810	0,282	85,741	4,000
65	shelf	0,01759	0,329	86,071	4,667
66	light	0,01709	0,188	86,259	2,667
67	laboratory	0,01457	0,282	86,541	4,000
68	netbook	0,01424	0,188	86,729	2,667
69	test	0,01409	0,188	86,918	2,667
70	Wall	0,01391	0,141	87,059	2,000
71	children	0,01364	0,188	87,247	2,667
72	Internet	0,01356	0,141	87,388	2,000
73	homework	0,01331	0,188	87,576	2,667
74	mark	0,01309	0,188	87,765	2,667
75	crayón	0,01284	0,188	87,953	2,667
76	radio	0,01275	0,188	88,141	2,667
77	laptop	0,01266	0,188	88,329	2,667
78	stair	0,01264	0,141	88,471	2,000
79	gymnastics	0,01262	0,141	88,612	2,000
80	exercise	0,01246	0,188	88,800	2,667
81	paint	0,01235	0,188	88,988	2,667
82	lesson	0,01229	0,141	89,129	2,000
83	backpack	0,01216	0,141	89,271	2,000
84	seat	0,01198	0,094	89,365	1,333
85	CD player	0,01139	0,141	89,506	2,000
86	suitcase	0,01122	0,141	89,647	2,000
87	scissors	0,01121	0,188	89,835	2,667
88	high school	0,01103	0,141	89,976	2,000
89	stapler	0,01075	0,188	90,165	2,667
90	speaker	0,01063	0,141	90,306	2,000
91	glue	0,01044	0,141	90,447	2,000
92	class list	0,01043	0,094	90,541	1,333
93	paper sheet	0,01043	0,094	90,635	1,333
94	Post-it	0,01034	0,141	90,776	2,000

95	PC	0,01009	0,141	90,918	2,000
96	study	0,01006	0,141	91,059	2,000
97	teenager	0,01003	0,094	91,153	1,333
98	hall	0,00933	0,094	91,247	1,333
99	closet	0,00927	0,094	91,341	1,333
100	mathematics	0,00903	0,094	91,435	1,333
101	English lesson	0,00899	0,094	91,529	1,333
102	highlighter	0,00888	0,188	91,718	2,667
103	file	0,00873	0,141	91,859	2,000
104	rucksack	0,00855	0,141	92,000	2,000
105	classmate	0,00821	0,094	92,094	1,333
106	bench	0,00819	0,094	92,188	1,333
107	armchair	0,00782	0,094	92,282	1,333
108	blind	0,00781	0,141	92,424	2,000
109	university	0,00778	0,094	92,518	1,333
110	draw	0,00757	0,141	92,659	2,000
111	history	0,00752	0,094	92,753	1,333
112	digital board	0,00727	0,094	92,847	1,333
113	headmaster	0,00715	0,094	92,941	1,333
114	picture	0,00685	0,094	93,035	1,333
115	kindergarten	0,00684	0,094	93,129	1,333
116	keyboard	0,00683	0,094	93,224	1,333
117	video	0,00678	0,094	93,318	1,333
118	copy	0,00674	0,094	93,412	1,333
119	furniture	0,00667	0,047	93,459	0,667
120	school	0,00667	0,047	93,506	0,667
121	friend	0,00665	0,094	93,600	1,333
122	task	0,00665	0,094	93,694	1,333
123	glasses	0,00644	0,094	93,788	1,333
124	pencil sharpener	0,00631	0,094	93,882	1,333
125	mirror	0,00620	0,047	93,929	0,667
126	ball	0,00606	0,094	94,024	1,333
127	floor	0,00606	0,094	94,118	1,333
128	learn	0,00600	0,094	94,212	1,333
129	adolescent	0,00577	0,047	94,259	0,667
130	cheat	0,00577	0,047	94,306	0,667
131	lift	0,00577	0,047	94,353	0,667
132	uncomfortable	0,00577	0,047	94,400	0,667
133	dull	0,00537	0,047	94,447	0,667
134	grammar	0,00537	0,047	94,494	0,667
135	music	0,00537	0,047	94,541	0,667
136	notice board	0,00537	0,047	94,588	0,667
137	vending machine	0,00537	0,047	94,635	0,667
138	ceiling	0,00513	0,094	94,729	1,333
139	principal	0,00508	0,094	94,824	1,333
140	boot	0,00500	0,047	94,871	0,667
141	desktop computer	0,00500	0,047	94,918	0,667
142	pavement	0,00500	0,047	94,965	0,667
143	piece of paper	0,00500	0,047	95,012	0,667

144	WC	0,00500	0,047	95,059	0,667
145	heat	0,00487	0,094	95,153	1,333
146	dean	0,00465	0,047	95,200	0,667
147	drawing	0,00465	0,047	95,247	0,667
148	garden	0,00465	0,047	95,294	0,667
149	light bulb	0,00465	0,047	95,341	0,667
150	master	0,00465	0,047	95,388	0,667
151	portfolio	0,00465	0,047	95,435	0,667
152	shirt	0,00465	0,047	95,482	0,667
153	text	0,00465	0,047	95,529	0,667
154	game	0,00444	0,094	95,624	1,333
155	agenda	0,00440	0,094	95,718	1,333
156	coursebook	0,00433	0,047	95,765	0,667
157	faculty	0,00433	0,047	95,812	0,667
158	leaf	0,00433	0,047	95,859	0,667
159	overhead projector	0,00433	0,047	95,906	0,667
160	tennis	0,00433	0,047	95,953	0,667
161	Wood	0,00433	0,047	96,000	0,667
162	physics	0,00405	0,094	96,094	1,333
163	bell	0,00403	0,047	96,141	0,667
164	hi-fi	0,00403	0,047	96,188	0,667
165	sandwich	0,00403	0,047	96,235	0,667
166	screwdriver	0,00403	0,047	96,282	0,667
167	bookshelf	0,00375	0,047	96,329	0,667
168	box	0,00375	0,047	96,376	0,667
169	dust cover	0,00375	0,047	96,424	0,667
170	English	0,00375	0,047	96,471	0,667
171	flashcard	0,00375	0,047	96,518	0,667
172	network	0,00375	0,047	96,565	0,667
173	panel	0,00375	0,047	96,612	0,667
174	partner	0,00375	0,047	96,659	0,667
175	people	0,00375	0,047	96,706	0,667
176	philosophy	0,00375	0,047	96,753	0,667
177	DVD	0,00354	0,094	96,847	1,333
178	blog	0,00349	0,047	96,894	0,667
179	cassette	0,00349	0,047	96,941	0,667
180	encyclopedia	0,00349	0,047	96,988	0,667
181	football match	0,00349	0,047	97,035	0,667
182	microphone	0,00349	0,047	97,082	0,667
183	sofá	0,00349	0,047	97,129	0,667
184	biology	0,00325	0,047	97,176	0,667
185	clip	0,00325	0,047	97,224	0,667
186	master class	0,00325	0,047	97,271	0,667
187	roof	0,00325	0,047	97,318	0,667
188	science	0,00325	0,047	97,365	0,667
189	shoulder bag	0,00325	0,047	97,412	0,667
190	stereo	0,00325	0,047	97,459	0,667
191	bed	0,00302	0,047	97,506	0,667
192	bookcase	0,00302	0,047	97,553	0,667

193	correcting fluid	0,00302	0,047	97,600	0,667
194	cutter	0,00302	0,047	97,647	0,667
195	DVD player	0,00302	0,047	97,694	0,667
196	entrance hall	0,00302	0,047	97,741	0,667
197	Internet access	0,00302	0,047	97,788	0,667
198	personal computer	0,00302	0,047	97,835	0,667
199	primary school	0,00302	0,047	97,882	0,667
200	punishment	0,00302	0,047	97,929	0,667
201	racket	0,00302	0,047	97,976	0,667
202	sheet of paper	0,00302	0,047	98,024	0,667
203	sponge	0,00302	0,047	98,071	0,667
204	cable	0,00281	0,047	98,118	0,667
205	office	0,00281	0,047	98,165	0,667
206	pin	0,00281	0,047	98,212	0,667
207	compass	0,00262	0,047	98,259	0,667
208	concierge	0,00262	0,047	98,306	0,667
209	cork	0,00262	0,047	98,353	0,667
210	knowledge	0,00262	0,047	98,400	0,667
211	plug	0,00262	0,047	98,447	0,667
212	dining hall	0,00244	0,047	98,494	0,667
213	newspaper	0,00244	0,047	98,541	0,667
214	write	0,00244	0,047	98,588	0,667
215	cafetería	0,00227	0,047	98,635	0,667
216	read	0,00227	0,047	98,682	0,667
217	pool	0,00211	0,047	98,729	0,667
218	sticker	0,00211	0,047	98,776	0,667
219	case	0,00196	0,047	98,824	0,667
220	clock	0,00196	0,047	98,871	0,667
221	corridor	0,00196	0,047	98,918	0,667
222	flute	0,00196	0,047	98,965	0,667
223	litter	0,00196	0,047	99,012	0,667
224	Plasticine	0,00196	0,047	99,059	0,667
225	art college	0,00183	0,047	99,106	0,667
226	coulored pencil	0,00183	0,047	99,153	0,667
227	diary	0,00183	0,047	99,200	0,667
228	essay	0,00183	0,047	99,247	0,667
229	Play-Doh	0,00183	0,047	99,294	0,667
230	assessment	0,00170	0,047	99,341	0,667
231	pass	0,00170	0,047	99,388	0,667
232	plant	0,00170	0,047	99,435	0,667
233	rubbish	0,00170	0,047	99,482	0,667
234	Sellotape	0,00170	0,047	99,529	0,667
235	deadline	0,00158	0,047	99,576	0,667
236	fail	0,00158	0,047	99,624	0,667
237	parking	0,00158	0,047	99,671	0,667
238	puncher	0,00158	0,047	99,718	0,667
239	course	0,00147	0,047	99,765	0,667
240	oral examination	0,00147	0,047	99,812	0,667
241	waste paper	0,00147	0,047	99,859	0,667

242	written examination	0,00137	0,047	99,906	0,667
243	staple	0,00128	0,047	99,953	0,667
244	language	0,00089	0,047	100,000	0,667

2.2.5. The city

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	car	0,54884	4,561	4,561	71,333
2	Street	0,50812	4,177	8,738	65,333
3	building	0,43831	3,623	12,361	56,667
4	park	0,29681	3,069	15,431	48,000
5	bus	0,29564	2,728	18,159	42,667
6	people	0,27362	2,856	21,014	44,667
7	shop	0,26184	2,813	23,828	44,000
8	house	0,24468	2,344	26,172	36,667
9	road	0,21935	2,003	28,176	31,333
10	traffic light	0,17243	1,833	30,009	28,667
11	square	0,16913	1,662	31,671	26,000
12	restaurant	0,15509	1,705	33,376	26,667
13	tree	0,14805	1,620	34,996	25,333
14	school	0,13856	1,577	36,573	24,667
15	museum	0,13699	1,492	38,065	23,333
16	supermarket	0,13360	1,577	39,642	24,667
17	cinema	0,12876	1,449	41,091	22,667
18	underground	0,11573	1,108	42,199	17,333
19	flat	0,11041	1,108	43,308	17,333
20	church	0,11037	1,066	44,373	16,667
21	taxi	0,10067	1,194	45,567	18,667
22	bank	0,09605	1,108	46,675	17,333
23	shopping centre	0,09356	1,108	47,783	17,333
24	motorbike	0,09110	0,895	48,679	14,000
25	university	0,09041	1,194	49,872	18,667
26	traffic	0,09024	0,980	50,853	15,333
27	monument	0,08942	0,853	51,705	13,333
28	train	0,08700	0,895	52,600	14,000
29	theatre	0,08494	1,023	53,623	16,000
30	avenue	0,08465	0,895	54,518	14,000
31	pollution	0,08379	0,980	55,499	15,333
32	bar	0,08270	0,938	56,436	14,667
33	bus stop	0,08059	1,066	57,502	16,667
34	hospital	0,08028	0,853	58,355	13,333
35	traffic jam	0,08003	0,895	59,250	14,000
36	pub	0,07992	0,895	60,145	14,000
37	garden	0,07590	0,895	61,040	14,000
38	airport	0,07504	0,938	61,978	14,667
39	skyscraper	0,06940	0,639	62,617	10,000
40	noise	0,06808	0,682	63,299	10,667
41	train station	0,06743	0,853	64,152	13,333
42	town hall	0,06724	0,682	64,834	10,667
43	parking	0,06691	0,767	65,601	12,000
44	light	0,06585	0,554	66,155	8,667
45	city centre	0,06255	0,639	66,795	10,000

46	library	0,06166	0,767	67,562	12,000
47	pólice	0,05961	0,682	68,244	10,667
48	bus station	0,05945	0,639	68,883	10,000
49	cathedral	0,05555	0,639	69,523	10,000
50	pedestrian	0,05325	0,639	70,162	10,000
51	neighbourhood	0,05150	0,469	70,631	7,333
52	mal	0,05079	0,767	71,398	12,000
53	city hall	0,04904	0,512	71,910	8,000
54	bicycle	0,04658	0,512	72,421	8,000
55	roundabout	0,04563	0,426	72,847	6,667
56	bridge	0,04454	0,512	73,359	8,000
57	market	0,04407	0,597	73,956	9,333
58	pavement	0,04107	0,384	74,339	6,000
59	bike	0,03900	0,384	74,723	6,000
60	office	0,03711	0,384	75,107	6,000
61	fountain	0,03503	0,384	75,490	6,000
62	dog	0,03261	0,384	75,874	6,000
63	metro	0,03173	0,298	76,172	4,667
64	tube	0,03173	0,341	76,513	5,333
65	centre	0,03173	0,298	76,812	4,667
66	hotel	0,03070	0,426	77,238	6,667
67	post office	0,03011	0,298	77,536	4,667
68	police station	0,02957	0,384	77,920	6,000
69	beach	0,02869	0,298	78,218	4,667
70	port	0,02826	0,384	78,602	6,000
71	commercial centre	0,02755	0,341	78,943	5,333
72	castle	0,02607	0,213	79,156	3,333
73	disco(theque)	0,02581	0,256	79,412	4,000
74	subway	0,02528	0,298	79,710	4,667
75	policeman	0,02486	0,341	80,051	5,333
76	place	0,02445	0,171	80,222	2,667
77	statue	0,02398	0,256	80,477	4,000
78	highway	0,02343	0,298	80,776	4,667
79	river	0,02147	0,298	81,074	4,667
80	promenade	0,02104	0,213	81,287	3,333
81	bookshop	0,02062	0,256	81,543	4,000
82	store	0,02031	0,213	81,756	3,333
83	Málaga	0,01924	0,171	81,927	2,667
84	rubbish	0,01892	0,213	82,140	3,333
85	citizen	0,01891	0,213	82,353	3,333
86	London	0,01860	0,171	82,523	2,667
87	stress	0,01848	0,213	82,737	3,333
88	signal	0,01837	0,213	82,950	3,333
89	council	0,01690	0,213	83,163	3,333
90	neighbour	0,01587	0,171	83,333	2,667
91	high school	0,01581	0,171	83,504	2,667
92	children	0,01560	0,256	83,760	4,000
93	cat	0,01517	0,171	83,930	2,667
94	bench	0,01495	0,213	84,143	3,333

95	cafeteria	0,01463	0,171	84,314	2,667
96	boulevard	0,01455	0,171	84,484	2,667
97	mountain	0,01431	0,128	84,612	2,000
98	person	0,01386	0,128	84,740	2,000
99	butcher's shop	0,01383	0,171	84,910	2,667
100	New York	0,01377	0,171	85,081	2,667
101	motorway	0,01361	0,171	85,251	2,667
102	chemist's shop	0,01352	0,128	85,379	2,000
103	swimming pool	0,01350	0,171	85,550	2,667
104	zebra crossing	0,01298	0,171	85,720	2,667
105	car park	0,01282	0,171	85,891	2,667
106	café	0,01274	0,171	86,061	2,667
107	motorcycle	0,01249	0,171	86,232	2,667
108	suburb	0,01239	0,213	86,445	3,333
109	citadel	0,01209	0,128	86,573	2,000
110	sky	0,01198	0,085	86,658	1,333
111	harbour	0,01190	0,128	86,786	2,000
112	circus	0,01167	0,085	86,871	1,333
113	lift	0,01155	0,085	86,957	1,333
114	outskirts	0,01143	0,128	87,084	2,000
115	station	0,01143	0,128	87,212	2,000
116	plane	0,01121	0,085	87,298	1,333
117	map	0,01102	0,128	87,425	2,000
118	metropolitan	0,01077	0,085	87,511	1,333
119	town	0,01072	0,128	87,639	2,000
120	sea	0,01057	0,128	87,766	2,000
121	downtown	0,01008	0,128	87,894	2,000
122	apartment	0,01003	0,085	87,980	1,333
123	gym	0,00988	0,171	88,150	2,667
124	stop	0,00970	0,085	88,235	1,333
125	semaphore	0,00965	0,085	88,321	1,333
126	corruption	0,00929	0,085	88,406	1,333
127	skyline	0,00927	0,085	88,491	1,333
128	cab	0,00926	0,128	88,619	2,000
129	shopping mall	0,00918	0,128	88,747	2,000
130	van	0,00902	0,085	88,832	1,333
131	club	0,00902	0,085	88,917	1,333
132	window	0,00875	0,085	89,003	1,333
133	village	0,00863	0,128	89,130	2,000
134	playground	0,00849	0,085	89,216	1,333
135	pedestrian crossing	0,00841	0,085	89,301	1,333
136	plant	0,00828	0,128	89,429	2,000
137	concert	0,00819	0,085	89,514	1,333
138	seafront	0,00808	0,085	89,599	1,333
139	musical	0,00802	0,085	89,685	1,333
140	contamination	0,00778	0,085	89,770	1,333
141	stadium	0,00768	0,085	89,855	1,333
142	nightclub	0,00768	0,085	89,940	1,333
143	friend	0,00698	0,085	90,026	1,333

144	lamp	0,00698	0,085	90,111	1,333
145	fortress	0,00695	0,085	90,196	1,333
146	laundry	0,00695	0,085	90,281	1,333
147	problem	0,00684	0,085	90,367	1,333
148	tower	0,00678	0,085	90,452	1,333
149	lake	0,00670	0,085	90,537	1,333
150	bump	0,00667	0,043	90,580	0,667
151	capital city	0,00667	0,043	90,622	0,667
152	city	0,00667	0,043	90,665	0,667
153	landscape	0,00667	0,043	90,708	0,667
154	major	0,00667	0,043	90,750	0,667
155	Oxford	0,00667	0,043	90,793	0,667
156	round	0,00667	0,043	90,835	0,667
157	sex	0,00667	0,043	90,878	0,667
158	sightseeing	0,00667	0,043	90,921	0,667
159	Torremolinos	0,00667	0,043	90,963	0,667
160	fireman	0,00649	0,128	91,091	2,000
161	ambulance	0,00637	0,085	91,176	1,333
162	zoo	0,00636	0,128	91,304	2,000
163	flower	0,00634	0,128	91,432	2,000
164	advertising	0,00631	0,085	91,517	1,333
165	Benalmádena	0,00620	0,043	91,560	0,667
166	company	0,00620	0,043	91,603	0,667
167	home	0,00620	0,043	91,645	0,667
168	train stop	0,00620	0,043	91,688	0,667
169	travel	0,00620	0,043	91,731	0,667
170	work	0,00603	0,085	91,816	1,333
171	homelessness	0,00586	0,085	91,901	1,333
172	litter	0,00584	0,085	91,986	1,333
173	grass	0,00577	0,043	92,029	0,667
174	Mijas	0,00577	0,043	92,072	0,667
175	near	0,00577	0,043	92,114	0,667
176	radio	0,00577	0,043	92,157	0,667
177	roof	0,00577	0,043	92,199	0,667
178	smog	0,00577	0,043	92,242	0,667
179	smoke	0,00577	0,043	92,285	0,667
180	crime	0,00563	0,085	92,370	1,333
181	shopping	0,00558	0,085	92,455	1,333
182	crowd	0,00545	0,085	92,540	1,333
183	hairdresser	0,00540	0,085	92,626	1,333
184	accident	0,00537	0,043	92,668	0,667
185	billboard	0,00537	0,043	92,711	0,667
186	dirt	0,00537	0,043	92,754	0,667
187	earth	0,00537	0,043	92,796	0,667
188	electricity	0,00537	0,043	92,839	0,667
189	Fuengirola	0,00537	0,043	92,882	0,667
190	ground	0,00537	0,043	92,924	0,667
191	path	0,00537	0,043	92,967	0,667
192	sidewalk	0,00529	0,085	93,052	1,333

193	mailbox	0,00525	0,085	93,137	1,333
194	child	0,00500	0,043	93,180	0,667
195	Marbella	0,00500	0,043	93,223	0,667
196	tram	0,00500	0,043	93,265	0,667
197	truck	0,00500	0,043	93,308	0,667
198	tube station	0,00500	0,043	93,350	0,667
199	litter bin	0,00489	0,085	93,436	1,333
200	obelisk	0,00489	0,085	93,521	1,333
201	money	0,00477	0,085	93,606	1,333
202	diversion	0,00471	0,085	93,691	1,333
203	stair	0,00471	0,085	93,777	1,333
204	airplane	0,00465	0,043	93,819	0,667
205	bin	0,00465	0,043	93,862	0,667
206	crew	0,00465	0,043	93,905	0,667
207	graveyard	0,00465	0,043	93,947	0,667
208	hill	0,00465	0,043	93,990	0,667
209	main way	0,00465	0,043	94,032	0,667
210	quarter	0,00465	0,043	94,075	0,667
211	show	0,00465	0,043	94,118	0,667
212	speed	0,00465	0,043	94,160	0,667
213	phone box	0,00455	0,085	94,246	1,333
214	newspaper	0,00454	0,085	94,331	1,333
215	casino	0,00433	0,043	94,373	0,667
216	cycle lane	0,00433	0,043	94,416	0,667
217	danger	0,00433	0,043	94,459	0,667
218	door	0,00433	0,043	94,501	0,667
219	fire station	0,00433	0,043	94,544	0,667
220	intersection	0,00433	0,043	94,587	0,667
221	radar	0,00433	0,043	94,629	0,667
222	snack bar	0,00433	0,043	94,672	0,667
223	town centre	0,00433	0,043	94,714	0,667
224	alley	0,00403	0,043	94,757	0,667
225	bingo	0,00403	0,043	94,800	0,667
226	coffee	0,00403	0,043	94,842	0,667
227	fish shop	0,00403	0,043	94,885	0,667
228	helicopter	0,00403	0,043	94,928	0,667
229	main street	0,00403	0,043	94,970	0,667
230	seaside	0,00403	0,043	95,013	0,667
231	vehicle	0,00403	0,043	95,055	0,667
232	public transport	0,00392	0,085	95,141	1,333
233	hall	0,00379	0,085	95,226	1,333
234	animal	0,00375	0,043	95,269	0,667
235	discussion	0,00375	0,043	95,311	0,667
236	Dublin	0,00375	0,043	95,354	0,667
237	flower shop	0,00375	0,043	95,396	0,667
238	Irish pub	0,00375	0,043	95,439	0,667
239	job	0,00375	0,043	95,482	0,667
240	mail	0,00375	0,043	95,524	0,667
241	main road	0,00375	0,043	95,567	0,667

242	tea	0,00375	0,043	95,610	0,667
243	gypsy	0,00358	0,085	95,695	1,333
244	coffee shop	0,00349	0,043	95,737	0,667
245	hedge	0,00349	0,043	95,780	0,667
246	inhabitant	0,00349	0,043	95,823	0,667
247	institution	0,00349	0,043	95,865	0,667
248	lighthouse	0,00349	0,043	95,908	0,667
249	secondary school	0,00349	0,043	95,951	0,667
250	street lamp	0,00349	0,043	95,993	0,667
251	walking	0,00349	0,043	96,036	0,667
252	Wheel	0,00349	0,043	96,078	0,667
253	book store	0,00340	0,085	96,164	1,333
254	baker's shop	0,00325	0,043	96,206	0,667
255	chemistry laboratory	0,00325	0,043	96,249	0,667
256	department store	0,00325	0,043	96,292	0,667
257	England	0,00325	0,043	96,334	0,667
258	fence	0,00325	0,043	96,377	0,667
259	firefighter	0,00325	0,043	96,419	0,667
260	fishmonger	0,00325	0,043	96,462	0,667
261	floor	0,00325	0,043	96,505	0,667
262	language	0,00325	0,043	96,547	0,667
263	music	0,00325	0,043	96,590	0,667
264	party	0,00325	0,043	96,633	0,667
265	penthouse	0,00325	0,043	96,675	0,667
266	postbox	0,00325	0,043	96,718	0,667
267	scooter	0,00325	0,043	96,760	0,667
268	sewage	0,00325	0,043	96,803	0,667
269	tourist	0,00325	0,043	96,846	0,667
270	unemployed people	0,00325	0,043	96,888	0,667
271	way	0,00325	0,043	96,931	0,667
272	culture	0,00302	0,043	96,974	0,667
273	luxury	0,00302	0,043	97,016	0,667
274	rush hour	0,00302	0,043	97,059	0,667
275	ship	0,00302	0,043	97,101	0,667
276	street market	0,00302	0,043	97,144	0,667
277	wagon	0,00302	0,043	97,187	0,667
278	butchery	0,00281	0,043	97,229	0,667
279	corner	0,00281	0,043	97,272	0,667
280	double-decker bus	0,00281	0,043	97,315	0,667
281	facade	0,00281	0,043	97,357	0,667
282	hypermarket	0,00281	0,043	97,400	0,667
283	monolith	0,00281	0,043	97,442	0,667
284	old	0,00281	0,043	97,485	0,667
285	parliament building	0,00281	0,043	97,528	0,667
286	petrol station	0,00281	0,043	97,570	0,667
287	poverty	0,00281	0,043	97,613	0,667

288	bird	0,00262	0,043	97,656	0,667
289	garage	0,00262	0,043	97,698	0,667
290	joy	0,00262	0,043	97,741	0,667
291	man	0,00262	0,043	97,783	0,667
292	mayor	0,00262	0,043	97,826	0,667
293	opera house	0,00262	0,043	97,869	0,667
294	railway station	0,00262	0,043	97,911	0,667
295	sightsee	0,00262	0,043	97,954	0,667
296	tea room	0,00262	0,043	97,997	0,667
297	tramway	0,00262	0,043	98,039	0,667
298	asphalt	0,00244	0,043	98,082	0,667
299	cottage	0,00244	0,043	98,124	0,667
300	dustbin	0,00244	0,043	98,167	0,667
301	government	0,00244	0,043	98,210	0,667
302	junction	0,00244	0,043	98,252	0,667
303	post	0,00244	0,043	98,295	0,667
304	woman	0,00244	0,043	98,338	0,667
305	hostel	0,00227	0,043	98,380	0,667
306	road sign	0,00227	0,043	98,423	0,667
307	semi-detached house	0,00227	0,043	98,465	0,667
308	stressed people	0,00227	0,043	98,508	0,667
309	taxicab	0,00227	0,043	98,551	0,667
310	villa	0,00227	0,043	98,593	0,667
311	clinic	0,00211	0,043	98,636	0,667
312	detached house	0,00211	0,043	98,679	0,667
313	kiosk	0,00211	0,043	98,721	0,667
314	market place	0,00211	0,043	98,764	0,667
315	musical comedy	0,00211	0,043	98,806	0,667
316	public works	0,00211	0,043	98,849	0,667
317	the Big Apple	0,00211	0,043	98,892	0,667
318	block	0,00196	0,043	98,934	0,667
319	fashion	0,00196	0,043	98,977	0,667
320	high street	0,00196	0,043	99,020	0,667
321	lamppost	0,00196	0,043	99,062	0,667
322	mansion	0,00196	0,043	99,105	0,667
323	pedestrian area	0,00196	0,043	99,147	0,667
324	tunnel	0,00196	0,043	99,190	0,667
325	main town	0,00183	0,043	99,233	0,667
326	stall	0,00183	0,043	99,275	0,667
327	campus	0,00170	0,043	99,318	0,667
328	old town	0,00170	0,043	99,361	0,667
329	barber's shop	0,00158	0,043	99,403	0,667
330	grocery store	0,00158	0,043	99,446	0,667
331	Paris	0,00158	0,043	99,488	0,667
332	railway	0,00158	0,043	99,531	0,667

333	smell	0,00158	0,043	99,574	0,667
334	underground works	0,00158	0,043	99,616	0,667
335	carpenter	0,00147	0,043	99,659	0,667
336	cemetery	0,00147	0,043	99,702	0,667
337	hurry	0,00147	0,043	99,744	0,667
338	Madrid	0,00147	0,043	99,787	0,667
339	sex shop	0,00147	0,043	99,829	0,667
340	clothes shop	0,00137	0,043	99,872	0,667
341	spring	0,00137	0,043	99,915	0,667
342	meadow	0,00128	0,043	99,957	0,667
343	history	0,00111	0,043	100,000	0,667

2.2.6. Means of transport

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	car	0,86363	8,716	8,716	97,333
2	bus	0,75300	7,761	16,478	86,667
3	train	0,65866	7,224	23,701	80,667
4	plane	0,49180	5,970	29,672	66,667
5	motorbike	0,40085	5,134	34,806	57,333
6	bike	0,39361	4,776	39,582	53,333
7	boat	0,34843	4,896	44,478	54,667
8	taxi	0,31296	3,940	48,418	44,000
9	bicycle	0,30400	3,701	52,119	41,333
10	ship	0,26889	3,761	55,881	42,000
11	underground	0,25290	3,224	59,104	36,000
12	helicopter	0,16129	2,567	61,672	28,667
13	horse	0,14798	2,269	63,940	25,333
14	airplane	0,14177	1,791	65,731	20,000
15	on foot	0,13473	1,970	67,701	22,000
16	metro	0,12108	1,552	69,254	17,333
17	motorcycle	0,11398	1,493	70,746	16,667
18	tube	0,11093	1,493	72,239	16,667
19	skate	0,09685	1,552	73,791	17,333
20	tram	0,08783	1,313	75,104	14,667
21	Lorry	0,08143	1,134	76,239	12,667
22	Truck	0,07550	1,254	77,493	14,000
23	yacht	0,07141	1,134	78,627	12,667
24	Cab	0,05939	0,836	79,463	9,333
25	aeroplane	0,05746	0,716	80,179	8,000
26	donkey	0,05662	0,896	81,075	10,000
27	subway	0,05305	0,716	81,791	8,000
28	Ferry	0,04601	0,776	82,567	8,667
29	Cycle	0,03909	0,478	83,045	5,333
30	cruiser	0,03846	0,597	83,642	6,667
31	submarine	0,03582	0,537	84,179	6,000
32	Van	0,03451	0,537	84,716	6,000
33	railway	0,03342	0,418	85,134	4,667
34	Jet	0,03241	0,478	85,612	5,333
35	skateboard	0,03063	0,418	86,030	4,667
36	roller skate	0,02825	0,478	86,507	5,333
37	cruise	0,02637	0,478	86,985	5,333
38	camel	0,02585	0,358	87,343	4,000
39	transatlantic	0,02125	0,299	87,642	3,333
40	scooter	0,02020	0,299	87,940	3,333
41	coach	0,02019	0,299	88,239	3,333
42	balloon	0,01790	0,299	88,537	3,333
43	airport	0,01784	0,299	88,836	3,333
44	Ski	0,01745	0,299	89,134	3,333
45	tricycle	0,01737	0,239	89,373	2,667

46	canoe	0,01683	0,299	89,672	3,333
47	trolley	0,01657	0,239	89,910	2,667
48	parachute	0,01580	0,358	90,269	4,000
49	caravan	0,01544	0,239	90,507	2,667
50	ambulance	0,01400	0,239	90,746	2,667
51	mountain bike	0,01335	0,179	90,925	2,000
52	rocket	0,01332	0,179	91,104	2,000
53	fast train	0,01293	0,239	91,343	2,667
54	spaceship	0,01293	0,239	91,582	2,667
55	walking frame	0,01150	0,179	91,761	2,000
56	elephant	0,01111	0,179	91,940	2,000
57	Leg	0,01086	0,119	92,060	1,333
58	aircraft	0,01037	0,119	92,179	1,333
59	double-decker bus	0,01037	0,119	92,299	1,333
60	Fly	0,01003	0,119	92,418	1,333
61	train station	0,00977	0,179	92,597	2,000
62	high-speed train	0,00933	0,119	92,716	1,333
63	ocean liner	0,00933	0,119	92,836	1,333
64	Shoe	0,00918	0,179	93,015	2,000
65	horse-drawn carriage	0,00903	0,119	93,134	1,333
66	tráiler	0,00875	0,119	93,254	1,333
67	Walk	0,00841	0,119	93,373	1,333
68	shuttle	0,00839	0,119	93,493	1,333
69	tramway	0,00750	0,119	93,612	1,333
70	paraglider	0,00705	0,119	93,731	1,333
71	animal	0,00697	0,119	93,851	1,333
72	Rollerblade	0,00684	0,119	93,970	1,333
73	ticket	0,00660	0,119	94,090	1,333
74	jet ski	0,00651	0,119	94,209	1,333
75	Road	0,00637	0,119	94,328	1,333
76	Sky	0,00637	0,119	94,448	1,333
77	kayak	0,00630	0,119	94,567	1,333
78	cyclist	0,00620	0,060	94,627	0,667
79	lifeboat	0,00620	0,060	94,687	0,667
80	sheep	0,00620	0,060	94,746	0,667
81	bus stop	0,00606	0,119	94,866	1,333
82	carriage	0,00606	0,119	94,985	1,333
83	limousine	0,00586	0,119	95,104	1,333
84	moped	0,00577	0,060	95,164	0,667
85	move	0,00577	0,060	95,224	0,667
86	sledge	0,00577	0,060	95,284	0,667
87	Hurry	0,00537	0,060	95,343	0,667
88	vehicle	0,00537	0,060	95,403	0,667
89	boat train	0,00500	0,060	95,463	0,667
90	flight	0,00500	0,060	95,522	0,667
91	rental car	0,00500	0,060	95,582	0,667
92	Trike	0,00500	0,060	95,642	0,667
93	gondola	0,00465	0,060	95,701	0,667

94	monorail	0,00465	0,060	95,761	0,667
95	travel	0,00465	0,060	95,821	0,667
96	caboose	0,00433	0,060	95,881	0,667
97	camping van	0,00433	0,060	95,940	0,667
98	hitchhike	0,00433	0,060	96,000	0,667
99	Jeep	0,00433	0,060	96,060	0,667
100	journey	0,00433	0,060	96,119	0,667
101	motor launch	0,00433	0,060	96,179	0,667
102	space shuttle	0,00433	0,060	96,239	0,667
103	tractor	0,00433	0,060	96,299	0,667
104	Zeppelin	0,00433	0,060	96,358	0,667
105	harvester	0,00403	0,060	96,418	0,667
106	hovercraft	0,00403	0,060	96,478	0,667
107	motorhome	0,00403	0,060	96,537	0,667
108	motorway	0,00403	0,060	96,597	0,667
109	Pony	0,00403	0,060	96,657	0,667
110	sailing boat	0,00403	0,060	96,716	0,667
111	Trip	0,00403	0,060	96,776	0,667
112	underway	0,00403	0,060	96,836	0,667
113	bulldozer	0,00375	0,060	96,896	0,667
114	metropolitan	0,00375	0,060	96,955	0,667
115	penny farthing	0,00375	0,060	97,015	0,667
116	racing yacht	0,00375	0,060	97,075	0,667
117	time machine	0,00375	0,060	97,134	0,667
118	trolleybus	0,00375	0,060	97,194	0,667
119	bus station	0,00349	0,060	97,254	0,667
120	driver	0,00349	0,060	97,313	0,667
121	highway	0,00349	0,060	97,373	0,667
122	police car	0,00349	0,060	97,433	0,667
123	Seat	0,00349	0,060	97,493	0,667
124	Tank	0,00349	0,060	97,552	0,667
125	bridge	0,00325	0,060	97,612	0,667
126	chariot	0,00325	0,060	97,672	0,667
127	dromedary	0,00325	0,060	97,731	0,667
128	jogging	0,00325	0,060	97,791	0,667
129	public transport	0,00325	0,060	97,851	0,667
130	shipment	0,00325	0,060	97,910	0,667
131	suburban rail	0,00325	0,060	97,970	0,667
132	Trans-Siberian Railway	0,00325	0,060	98,030	0,667
133	Cart	0,00302	0,060	98,090	0,667
134	elevator	0,00302	0,060	98,149	0,667
135	go home	0,00302	0,060	98,209	0,667
136	InterRail	0,00302	0,060	98,269	0,667
137	landing	0,00302	0,060	98,328	0,667
138	private transport	0,00302	0,060	98,388	0,667
139	Vespa	0,00302	0,060	98,448	0,667
140	wagon	0,00302	0,060	98,507	0,667
141	board	0,00281	0,060	98,567	0,667

142	Lift	0,00281	0,060	98,627	0,667
143	luxury transport	0,00281	0,060	98,687	0,667
144	scuba diving	0,00281	0,060	98,746	0,667
145	ski lift	0,00281	0,060	98,806	0,667
146	BMW	0,00262	0,060	98,866	0,667
147	cable car	0,00262	0,060	98,925	0,667
148	Duty	0,00262	0,060	98,985	0,667
149	stair lift	0,00262	0,060	99,045	0,667
150	wheelchair	0,00262	0,060	99,104	0,667
151	jet airliner	0,00244	0,060	99,164	0,667
152	queue	0,00244	0,060	99,224	0,667
153	teleport	0,00244	0,060	99,284	0,667
154	traffic lights	0,00244	0,060	99,343	0,667
155	Key	0,00227	0,060	99,403	0,667
156	sidecar	0,00227	0,060	99,463	0,667
157	Fare	0,00211	0,060	99,522	0,667
158	motor	0,00211	0,060	99,582	0,667
159	Oil	0,00211	0,060	99,642	0,667
160	oil station	0,00196	0,060	99,701	0,667
161	seaway	0,00196	0,060	99,761	0,667
162	City	0,00183	0,060	99,821	0,667
163	monocycle	0,00158	0,060	99,881	0,667
164	Titanic	0,00158	0,060	99,940	0,667
165	wáter	0,00147	0,060	100,000	0,667

2.2.7. Games and entertainments

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	aerobics	0,03835	0,523	6,623	60,234
2	aerodance	0,00225	0,052	0,662	54,386
3	American football	0,01031	0,105	1,325	49,708
4	amusement park	0,00242	0,052	0,662	50,292
5	appointment with friends	0,00260	0,052	0,662	40,936
6	arcade	0,00195	0,052	0,662	36,842
7	art	0,00869	0,105	1,325	35,088
8	art gallery	0,00373	0,052	0,662	35,673
9	athletics	0,00704	0,105	1,325	32,749
10	backgammon	0,02584	0,261	3,311	32,164
11	badminton	0,01654	0,209	2,649	32,749
12	ball	0,08918	0,889	11,258	30,409
13	ballet	0,01002	0,157	1,987	29,825
14	Bar	0,01964	0,314	3,974	27,485
15	baseball	0,09195	1,045	13,245	24,561
16	basketball	0,43288	4,443	56,291	23,392
17	beach	0,02107	0,314	3,974	21,637
18	bicycle	0,00534	0,052	0,662	19,298
19	bike	0,02149	0,209	2,649	20,468
20	billiard	0,00945	0,157	1,987	20,468
21	bingo	0,00347	0,052	0,662	16,959
22	blackjack	0,00813	0,105	1,325	18,713
23	blind man's buff	0,00797	0,105	1,325	17,544
24	Blog	0,00462	0,052	0,662	18,713
25	board game	0,01036	0,105	1,325	15,205
26	Book	0,03774	0,366	4,636	18,713
27	bowling	0,03471	0,470	5,960	17,544
28	boxing	0,00462	0,052	0,662	15,789
29	Breach	0,00242	0,052	0,662	15,789
30	bungee jumping	0,00462	0,052	0,662	14,035
31	Call of Duty	0,00662	0,052	0,662	12,865
32	camping	0,00776	0,105	1,325	13,450
33	Car	0,00300	0,052	0,662	11,696
34	Card	0,17485	1,934	24,503	11,111
35	card game	0,01147	0,105	1,325	12,281
36	Catch	0,00574	0,052	0,662	12,281
37	CD	0,00260	0,052	0,662	10,526
38	charade	0,00323	0,052	0,662	10,526
39	Chat	0,02071	0,314	3,974	10,526
40	chatter	0,00612	0,105	1,325	11,696
41	chess	0,18315	1,725	21,854	11,111
42	children	0,00759	0,105	1,325	9,942
43	Chinese food	0,00195	0,052	0,662	11,111
44	cinema	0,33948	3,868	49,007	10,526

45	circus	0,02269	0,261	3,311	10,526
46	Clean	0,00280	0,052	0,662	10,526
47	climbing	0,00809	0,105	1,325	9,357
48	Club	0,00210	0,052	0,662	9,357
49	clubbing	0,00323	0,052	0,662	8,772
50	Cluedo	0,00430	0,052	0,662	8,772
51	coach	0,00300	0,052	0,662	8,772
52	comic	0,00225	0,052	0,662	7,602
53	competition	0,00157	0,052	0,662	7,602
54	computer	0,07475	0,836	10,596	9,357
55	computer game	0,08028	0,836	10,596	8,187
56	Conan	0,00616	0,052	0,662	8,187
57	concert	0,08846	1,150	14,570	8,187
58	console	0,00616	0,052	0,662	6,433
59	cooking	0,02297	0,314	3,974	7,602
60	Crash Team Racing	0,00534	0,052	0,662	7,018
61	crawl	0,00323	0,052	0,662	7,018
62	cricket	0,01208	0,157	1,987	7,602
63	crossword	0,02934	0,366	4,636	7,018
64	cycling	0,05171	0,732	9,272	6,433
65	dance	0,14827	2,143	27,152	7,018
66	Dart	0,00400	0,052	0,662	7,018
67	Desperate Housewives	0,00225	0,052	0,662	5,263
68	Dice	0,03043	0,366	4,636	7,602
69	digital booklet	0,00574	0,052	0,662	5,848
70	disco(theque)	0,04015	0,575	7,285	6,433
71	diving	0,00547	0,105	1,325	6,433
72	Doll	0,03436	0,418	5,298	5,263
73	dominoes	0,01898	0,157	1,987	7,018
74	drawing	0,03115	0,418	5,298	4,678
75	Drink	0,01005	0,157	1,987	4,678
76	drinking bout	0,00260	0,052	0,662	5,263
77	drinking game	0,00672	0,105	1,325	5,263
78	Drive	0,00430	0,052	0,662	5,263
79	Drum	0,00996	0,105	1,325	4,678
80	DVD	0,00896	0,105	1,325	5,848
81	Eat	0,00300	0,052	0,662	5,263
82	eat out	0,00640	0,105	1,325	4,678
83	exercise	0,00260	0,052	0,662	5,263
84	extreme sport	0,00300	0,052	0,662	5,263
85	Facebook	0,02610	0,366	4,636	5,263
86	fashion designer	0,00400	0,052	0,662	5,263
87	festival	0,00972	0,157	1,987	5,848
88	Field	0,00400	0,052	0,662	5,263
89	Fight	0,00242	0,052	0,662	4,678
90	Film	0,07326	0,784	9,934	4,678
91	fishing	0,01479	0,261	3,311	4,094
92	fitness	0,00940	0,157	1,987	3,509

93	football	0,55680	5,332	67,550	4,094
94	football league	0,00323	0,052	0,662	4,094
95	football match	0,01297	0,157	1,987	4,094
96	Formula 1	0,00534	0,052	0,662	3,509
97	friend	0,00869	0,105	1,325	4,678
98	fun run	0,00210	0,052	0,662	3,509
99	Game	0,01190	0,105	1,325	3,509
100	Game Boy	0,00373	0,052	0,662	2,924
101	gin rummy	0,00400	0,052	0,662	4,094
102	go for a walk	0,01160	0,157	1,987	4,094
103	go out	0,04030	0,523	6,623	4,094
104	go out for a drink	0,00260	0,052	0,662	3,509
105	go out to dinner	0,00785	0,157	1,987	3,509
106	go out with friends	0,03895	0,680	8,609	3,509
107	go shopping	0,03253	0,418	5,298	2,924
108	Golf	0,05292	0,627	7,947	3,509
109	Greek food	0,00182	0,052	0,662	2,924
110	Green Farm	0,00462	0,052	0,662	3,509
111	Guess Who?	0,00373	0,052	0,662	3,509
112	guitar	0,00844	0,105	1,325	3,509
113	Guitar Hero	0,00373	0,052	0,662	2,924
114	Gym	0,02104	0,418	5,298	2,924
115	gymnastics	0,00777	0,105	1,325	2,924
116	handball	0,03398	0,418	5,298	2,924
117	hang out with somebody	0,00280	0,052	0,662	2,924
118	Hide	0,00347	0,052	0,662	2,339
119	hide-and-seek	0,08412	0,889	11,258	2,339
120	hiking	0,01264	0,157	1,987	2,924
121	hockey	0,04402	0,523	6,623	2,924
122	Horse	0,00462	0,052	0,662	2,924
123	horse racing	0,00497	0,052	0,662	2,339
124	horse riding	0,00896	0,105	1,325	2,924
125	How I Met Your Mother	0,00462	0,052	0,662	2,339
126	hunting	0,00347	0,052	0,662	2,924
127	ice hockey	0,00347	0,052	0,662	2,339
128	ice skating	0,01501	0,157	1,987	2,924
129	Internet	0,07342	0,993	12,583	2,924
130	jazz concert	0,00462	0,052	0,662	2,924
131	jogging	0,10001	1,411	17,881	2,339
132	Joke	0,00662	0,052	0,662	2,924
133	jumping	0,00505	0,105	1,325	2,924
134	Just Dance	0,00195	0,052	0,662	2,339
135	karaoke	0,04775	0,627	7,947	2,339
136	karting	0,00776	0,105	1,325	2,924
137	kayak	0,00657	0,105	1,325	2,924
138	keyboard	0,00574	0,052	0,662	2,339
139	Lake	0,00323	0,052	0,662	2,924

140	laptop	0,00497	0,052	0,662	2,339
141	Lost	0,00534	0,052	0,662	2,339
142	Ludo	0,02567	0,261	3,311	1,754
143	magazine	0,00347	0,052	0,662	2,339
144	match	0,00400	0,052	0,662	2,339
145	meeting	0,00323	0,052	0,662	2,339
146	merry-go-round	0,00662	0,052	0,662	2,339
147	Monkey Island	0,00616	0,052	0,662	2,924
148	Monopoly	0,05999	0,732	9,272	1,754
149	motor racing	0,00574	0,052	0,662	2,339
150	motorcycling	0,00985	0,105	1,325	2,339
151	mountain	0,00799	0,105	1,325	2,339
152	movie	0,06700	0,784	9,934	1,754
153	museum	0,01500	0,261	3,311	2,339
154	music	0,25773	3,136	39,735	2,339
155	musical comedy	0,01089	0,157	1,987	1,754
156	nature	0,00430	0,052	0,662	2,339
157	network	0,00462	0,052	0,662	1,754
158	newspaper	0,00373	0,052	0,662	1,754
159	nightclub	0,01107	0,105	1,325	1,754
160	Nintendo	0,00820	0,105	1,325	1,754
161	opera	0,03948	0,575	7,285	1,754
162	paddle	0,07405	0,889	11,258	2,339
163	Paint	0,00373	0,052	0,662	1,754
164	painting	0,03731	0,418	5,298	1,754
165	paragliding	0,00373	0,052	0,662	1,754
166	Parcheesi	0,02340	0,209	2,649	1,754
167	Park	0,00858	0,157	1,987	1,754
168	Party	0,01938	0,209	2,649	2,339
169	password	0,00528	0,105	1,325	1,754
170	personal computer	0,00462	0,052	0,662	1,754
171	pétanque	0,01555	0,157	1,987	1,754
172	Petz	0,00430	0,052	0,662	2,339
173	pone	0,00260	0,052	0,662	1,754
174	photography	0,01718	0,209	2,649	1,754
175	piano	0,00534	0,052	0,662	1,754
176	Pictionary	0,00534	0,052	0,662	1,754
177	Pilates	0,00373	0,052	0,662	1,754
178	pinball	0,00497	0,052	0,662	1,754
179	ping-pong	0,01231	0,157	1,987	1,170
180	Play	0,02442	0,261	3,311	1,754
181	play a musical instrument	0,02676	0,314	3,974	1,754
182	play music	0,01532	0,209	2,649	1,170
183	Play Station	0,13897	1,464	18,543	1,754
184	play the cello	0,00430	0,052	0,662	1,754
185	play the clarinet	0,00242	0,052	0,662	1,170
186	play the flute	0,00400	0,052	0,662	1,754
187	play the guitar	0,02381	0,261	3,311	1,754

188	play the piano	0,01279	0,209	2,649	1,754
189	playground	0,00574	0,052	0,662	1,754
190	Playmobil	0,00347	0,052	0,662	2,339
191	poker	0,07855	0,836	10,596	1,170
192	Polo	0,01477	0,157	1,987	1,754
193	Pub	0,03907	0,523	6,623	1,170
194	puzzle	0,01422	0,209	2,649	1,754
195	quest	0,00497	0,052	0,662	1,754
196	Rabbids	0,00497	0,052	0,662	1,170
197	racing	0,00497	0,052	0,662	1,754
198	racket	0,00964	0,105	1,325	1,170
199	Radio	0,05367	0,836	10,596	1,170
200	rafting	0,00462	0,052	0,662	1,170
201	Rally	0,00731	0,105	1,325	1,754
202	Read	0,00532	0,105	1,325	1,170
203	read a book	0,03406	0,418	5,298	1,170
204	Reading	0,13871	1,673	21,192	1,170
205	Relax	0,00574	0,052	0,662	1,170
206	restaurant	0,01582	0,209	2,649	1,170
207	ride a horse	0,01644	0,261	3,311	1,170
208	riding	0,01086	0,157	1,987	1,170
209	Risk	0,00400	0,052	0,662	1,170
210	role play	0,02802	0,261	3,311	1,754
211	roller skating	0,00863	0,105	1,325	1,170
212	Rollerblade	0,00906	0,105	1,325	1,170
213	roulette	0,00462	0,052	0,662	1,170
214	rugby	0,06212	0,732	9,272	1,170
215	rugby match	0,00534	0,052	0,662	1,170
216	running	0,10046	1,307	16,556	1,170
217	sailing	0,01002	0,157	1,987	1,170
218	sandbox	0,00400	0,052	0,662	1,170
219	Scrabble	0,02388	0,261	3,311	1,170
220	scuba diving	0,00430	0,052	0,662	1,170
221	sculpture	0,00300	0,052	0,662	1,754
222	Sea	0,00616	0,052	0,662	1,170
223	seesaw	0,00616	0,052	0,662	1,170
224	serial	0,00574	0,052	0,662	1,754
225	series	0,02332	0,261	3,311	1,170
226	Shop	0,00323	0,052	0,662	1,170
227	shopping centre	0,00300	0,052	0,662	1,170
228	Show	0,01413	0,157	1,987	1,170
229	sightseeing	0,00580	0,105	1,325	1,170
230	Simon says	0,00863	0,105	1,325	1,170
231	Sims	0,00868	0,157	1,987	1,170
232	Sing	0,02072	0,261	3,311	1,170
233	singing	0,06905	0,941	11,921	1,754
234	Skate	0,03849	0,523	6,623	1,170
235	skating	0,01431	0,157	1,987	1,170
236	Ski	0,01715	0,261	3,311	1,170

237	skiing	0,03843	0,470	5,960	1,170
238	Skype	0,01295	0,157	1,987	1,170
239	slalom	0,00497	0,052	0,662	1,170
240	Sleep	0,00400	0,052	0,662	1,754
241	snakes and ladders	0,00497	0,052	0,662	1,170
242	snorkel	0,00136	0,052	0,662	1,170
243	Snow	0,00534	0,052	0,662	1,170
244	snowboard	0,01586	0,209	2,649	1,170
245	soap opera	0,01538	0,157	1,987	1,170
246	soccer	0,06921	0,732	9,272	1,170
247	social network	0,00497	0,052	0,662	1,170
248	softball	0,00939	0,105	1,325	1,170
249	speak	0,00784	0,157	1,987	1,170
250	speak with friends	0,00260	0,052	0,662	1,170
251	Spinning	0,02257	0,366	4,636	0,585
252	Sport	0,18367	2,091	26,490	0,585
253	squash	0,00673	0,105	1,325	0,585
254	Starbucks	0,00323	0,052	0,662	1,170
255	Step	0,00843	0,157	1,987	1,170
256	strategy game	0,00497	0,052	0,662	0,585
257	Study	0,00242	0,052	0,662	0,585
258	sudoku	0,01347	0,157	1,987	0,585
259	sunbathing	0,00242	0,052	0,662	0,585
260	Super Mario Bros	0,01157	0,157	1,987	0,585
261	surf	0,01477	0,209	2,649	0,585
262	surf the Net	0,00831	0,105	1,325	0,585
263	surfing	0,01192	0,157	1,987	0,585
264	swim	0,01159	0,105	1,325	1,170
265	swimming	0,12406	1,673	21,192	1,170
266	swimming pool	0,01273	0,157	1,987	0,585
267	swing	0,00574	0,052	0,662	0,585
268	table football	0,00347	0,052	0,662	0,585
269	table game	0,02970	0,314	3,974	1,170
270	table tennis	0,01959	0,261	3,311	1,170
271	Taboo	0,02726	0,261	3,311	1,170
272	take a coffee	0,00347	0,052	0,662	1,170
273	take a nap	0,00347	0,052	0,662	0,585
274	take a picture	0,00400	0,052	0,662	0,585
275	talk	0,00598	0,105	1,325	0,585
276	talk with friends	0,00623	0,105	1,325	0,585
277	team	0,00280	0,052	0,662	0,585
278	television	0,24257	3,032	38,411	0,585
279	television broadcast	0,00616	0,052	0,662	0,585
280	television programme	0,00574	0,052	0,662	0,585
281	Tennis	0,33085	3,607	45,695	0,585
282	Texas hold'em	0,00260	0,052	0,662	0,585

283	The Legend of Kyrandia	0,00662	0,052	0,662	0,585
284	Theatre	0,18061	2,196	27,815	0,585
285	trading card game	0,00497	0,052	0,662	0,585
286	Training	0,00462	0,052	0,662	0,585
287	Travel	0,05407	0,732	9,272	0,585
288	Trekking	0,01530	0,209	2,649	0,585
289	Triathlon	0,00300	0,052	0,662	0,585
290	trick or treat	0,00534	0,052	0,662	1,170
291	Trivial	0,03052	0,314	3,974	1,170
292	Tuenti	0,01050	0,157	1,987	0,585
293	Twitter	0,00430	0,052	0,662	0,585
294	Video	0,01539	0,157	1,987	0,585
295	videogame	0,30921	3,032	38,411	0,585
296	Volleyball	0,12142	1,411	17,881	0,585
297	Walk	0,00347	0,052	0,662	0,585
298	Walking	0,05291	0,680	8,609	0,585
299	water polo	0,01850	0,209	2,649	0,585
300	Wii	0,05334	0,627	7,947	0,585
301	window-shop	0,00323	0,052	0,662	1,170
302	Windsurf	0,00869	0,105	1,325	0,585
303	wine and dine	0,00146	0,052	0,662	0,585
304	World of Warcraft	0,01159	0,105	1,325	0,585
305	Wrestling	0,00225	0,052	0,662	0,585
306	write a book	0,00462	0,052	0,662	0,585
307	Writing	0,03016	0,470	5,960	0,585
308	Xbox	0,02324	0,261	3,311	0,585
309	Yahtzee	0,00430	0,052	0,662	0,585
310	Yoga	0,01278	0,209	2,649	0,585

2.2.8. Professions and jobs

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	teacher	0,84771	7,409	7,409	95,333
2	doctor	0,52568	5,337	12,746	68,667
3	policeman	0,37846	4,041	16,788	52,000
4	nurse	0,26060	2,902	19,689	37,333
5	lawyer	0,25652	2,798	22,487	36,000
6	engineer	0,20241	2,228	24,715	28,667
7	driver	0,20009	2,332	27,047	30,000
8	waiter	0,18576	2,228	29,275	28,667
9	architect	0,17297	1,917	31,192	24,667
10	fireman	0,16695	2,021	33,212	26,000
11	actor	0,15542	2,073	35,285	26,667
12	professor	0,13917	1,554	36,839	20,000
13	plumber	0,13514	1,399	38,238	18,000
14	writer	0,12553	1,554	39,793	20,000
15	painter	0,11890	1,451	41,244	18,667
16	taxi driver	0,11802	1,399	42,642	18,000
17	Singer	0,11453	1,399	44,041	18,000
18	butcher	0,10108	1,140	45,181	14,667
19	shop assistant	0,10092	1,295	46,477	16,667
20	cook	0,09965	1,295	47,772	16,667
21	bus driver	0,08817	1,036	48,808	13,333
22	manager	0,08324	1,036	49,845	13,333
23	barman	0,08148	0,984	50,829	12,667
24	pilot	0,07856	0,984	51,813	12,667
25	mechanic	0,07804	0,933	52,746	12,000
26	football player	0,07722	0,881	53,627	11,333
27	journalist	0,07591	0,984	54,611	12,667
28	musician	0,07540	0,881	55,492	11,333
29	Baker	0,07489	0,984	56,477	12,667
30	artista	0,07375	0,984	57,461	12,667
31	dancer	0,07317	0,933	58,394	12,000
32	designer	0,07073	1,036	59,430	13,333
33	secretary	0,06764	0,881	60,311	11,333
34	translator	0,06690	0,829	61,140	10,667
35	psychologist	0,06467	0,725	61,865	9,333
36	politician	0,06152	0,829	62,694	10,667
37	babysitter	0,06110	0,881	63,575	11,333
38	student	0,06085	0,725	64,301	9,333
39	dentist	0,05686	0,674	64,974	8,667
40	farmer	0,05254	0,674	65,648	8,667
41	carpenter	0,05050	0,518	66,166	6,667
42	judge	0,05012	0,622	66,788	8,000
43	footballer	0,04916	0,622	67,409	8,000
44	chef	0,04842	0,622	68,031	8,000
45	cleaner	0,04757	0,674	68,705	8,667

46	economist	0,04707	0,725	69,430	9,333
47	housewife	0,04670	0,622	70,052	8,000
48	hairdresser	0,04548	0,518	70,570	6,667
49	electrician	0,04496	0,518	71,088	6,667
50	assistant	0,04394	0,518	71,606	6,667
51	presenter	0,04306	0,674	72,280	8,667
52	director	0,04247	0,570	72,850	7,333
53	firefighter	0,04147	0,518	73,368	6,667
54	salesman	0,04081	0,466	73,834	6,000
55	gardener	0,03926	0,518	74,352	6,667
56	interpreter	0,03706	0,466	74,819	6,000
57	postman	0,03465	0,415	75,233	5,333
58	receptionist	0,03293	0,466	75,699	6,000
59	businessman	0,03248	0,466	76,166	6,000
60	solicitor	0,02975	0,415	76,580	5,333
61	actress	0,02897	0,415	76,995	5,333
62	sportsman	0,02855	0,363	77,358	4,667
63	Clerk	0,02799	0,363	77,720	4,667
64	guitarist	0,02674	0,363	78,083	4,667
65	researcher	0,02583	0,311	78,394	4,000
66	fisherman	0,02550	0,311	78,705	4,000
67	chemist	0,02546	0,415	79,119	5,333
68	scientist	0,02421	0,363	79,482	4,667
69	violinist	0,02419	0,259	79,741	3,333
70	president	0,02296	0,311	80,052	4,000
71	bróker	0,02170	0,311	80,363	4,000
72	boss	0,02160	0,259	80,622	3,333
73	housekeeper	0,02074	0,311	80,933	4,000
74	conductor	0,02006	0,259	81,192	3,333
75	historian	0,01961	0,207	81,399	2,667
76	physicist	0,01929	0,207	81,606	2,667
77	coach	0,01870	0,311	81,917	4,000
78	mayor	0,01859	0,259	82,176	3,333
79	tourist guide	0,01829	0,259	82,435	3,333
80	waitress	0,01816	0,311	82,746	4,000
81	security guard	0,01767	0,207	82,953	2,667
82	geographer	0,01752	0,207	83,161	2,667
83	bullfighter	0,01738	0,259	83,420	3,333
84	seller	0,01732	0,259	83,679	3,333
85	tennis player	0,01694	0,207	83,886	2,667
86	chief communications officer	0,01654	0,207	84,093	2,667
87	guide	0,01618	0,207	84,301	2,667
88	builder	0,01603	0,259	84,560	3,333
89	pianist	0,01549	0,207	84,767	2,667
90	prostitute	0,01503	0,155	84,922	2,000
91	linguist	0,01494	0,207	85,130	2,667
92	soldier	0,01482	0,259	85,389	3,333

93	computer programmer	0,01425	0,155	85,544	2,000
94	worker	0,01417	0,155	85,699	2,000
95	flight attendant	0,01410	0,259	85,959	3,333
96	barrister	0,01373	0,155	86,114	2,000
97	photographer	0,01373	0,155	86,269	2,000
98	draughtsman	0,01371	0,207	86,477	2,667
99	developer	0,01333	0,104	86,580	1,333
100	dustman	0,01320	0,155	86,736	2,000
101	lorry driver	0,01315	0,155	86,891	2,000
102	nun	0,01307	0,259	87,150	3,333
103	tourist agent	0,01302	0,155	87,306	2,000
104	sailor	0,01287	0,104	87,409	1,333
105	mathematician	0,01282	0,155	87,565	2,000
106	travel agent	0,01236	0,155	87,720	2,000
107	accountant	0,01223	0,155	87,876	2,000
108	basketball player	0,01188	0,155	88,031	2,000
109	chauffeur	0,01185	0,155	88,187	2,000
110	minister	0,01127	0,155	88,342	2,000
111	producer	0,01118	0,155	88,497	2,000
112	banker	0,01095	0,259	88,756	3,333
113	piano player	0,01095	0,155	88,912	2,000
114	constructor	0,01060	0,155	89,067	2,000
115	graphic designer	0,01052	0,155	89,223	2,000
116	bartender	0,01012	0,155	89,378	2,000
117	steward	0,01000	0,104	89,482	1,333
118	bricklayer	0,00981	0,155	89,637	2,000
119	tailor	0,00953	0,104	89,741	1,333
120	fishmonger	0,00933	0,104	89,845	1,333
121	civil servant	0,00931	0,104	89,948	1,333
122	police officer	0,00902	0,104	90,052	1,333
123	bóxer	0,00880	0,104	90,155	1,333
124	locksmith	0,00880	0,104	90,259	1,333
125	headmaster	0,00840	0,104	90,363	1,333
126	nanny	0,00840	0,104	90,466	1,333
127	cashier	0,00836	0,104	90,570	1,333
128	workman	0,00819	0,104	90,674	1,333
129	lecturer	0,00778	0,155	90,829	2,000
130	poet	0,00733	0,155	90,984	2,000
131	sales manager	0,00728	0,104	91,088	1,333
132	doorman	0,00724	0,104	91,192	1,333
133	priest	0,00705	0,155	91,347	2,000
134	curator	0,00700	0,104	91,451	1,333
135	trainer	0,00700	0,104	91,554	1,333
136	veterinarian	0,00700	0,104	91,658	1,333
137	reporter	0,00674	0,104	91,762	1,333
138	construction engineer	0,00667	0,052	91,813	0,667
139	gravedigger	0,00667	0,052	91,865	0,667

140	nursery nurse	0,00667	0,052	91,917	0,667
141	software engineer	0,00667	0,052	91,969	0,667
142	personal assistant	0,00665	0,104	92,073	1,333
143	conciierge	0,00651	0,104	92,176	1,333
144	fighter	0,00651	0,104	92,280	1,333
145	employee	0,00631	0,104	92,383	1,333
146	military policeman	0,00627	0,104	92,487	1,333
147	advertiser	0,00620	0,052	92,539	0,667
148	chemical engineer	0,00620	0,052	92,591	0,667
149	employer	0,00620	0,052	92,642	0,667
150	pimp	0,00620	0,052	92,694	0,667
151	speaker	0,00620	0,052	92,746	0,667
152	surveyor	0,00620	0,052	92,798	0,667
153	webmaster	0,00620	0,052	92,850	0,667
154	au pair	0,00605	0,104	92,953	1,333
155	supermodel	0,00602	0,104	93,057	1,333
156	shopper	0,00584	0,104	93,161	1,333
157	clown	0,00580	0,104	93,264	1,333
158	agent	0,00577	0,052	93,316	0,667
159	cellist	0,00577	0,052	93,368	0,667
160	film director	0,00577	0,052	93,420	0,667
161	illustrator	0,00577	0,052	93,472	0,667
162	labourer	0,00577	0,052	93,523	0,667
163	news agent	0,00577	0,052	93,575	0,667
164	officer	0,00577	0,052	93,627	0,667
165	personal trainer	0,00577	0,052	93,679	0,667
166	robber	0,00577	0,052	93,731	0,667
167	soccer player	0,00577	0,052	93,782	0,667
168	stylist	0,00577	0,052	93,834	0,667
169	entrepreneur	0,00568	0,155	93,990	2,000
170	bank robber	0,00537	0,052	94,041	0,667
171	blacksmith	0,00537	0,052	94,093	0,667
172	commercial traveller	0,00537	0,052	94,145	0,667
173	electrical engineer	0,00537	0,052	94,197	0,667
174	film producer	0,00537	0,052	94,249	0,667
175	paramedic	0,00537	0,052	94,301	0,667
176	philosopher	0,00537	0,052	94,352	0,667
177	programmer	0,00537	0,052	94,404	0,667
178	stripper	0,00537	0,052	94,456	0,667
179	systems engineer	0,00537	0,052	94,508	0,667
180	tattooist	0,00537	0,052	94,560	0,667
181	freelancer	0,00536	0,104	94,663	1,333
182	personal shopper	0,00525	0,104	94,767	1,333
183	physician	0,00525	0,104	94,870	1,333
184	annalist	0,00500	0,052	94,922	0,667
185	attorney	0,00500	0,052	94,974	0,667
186	Butler	0,00500	0,052	95,026	0,667
187	coach driver	0,00500	0,052	95,078	0,667
188	explorer	0,00500	0,052	95,130	0,667

189	financial adviser	0,00500	0,052	95,181	0,667
190	forensic	0,00500	0,052	95,233	0,667
191	headmistress	0,00500	0,052	95,285	0,667
192	nutritionist	0,00500	0,052	95,337	0,667
193	office worker	0,00500	0,052	95,389	0,667
194	porn actress	0,00500	0,052	95,440	0,667
195	typist	0,00500	0,052	95,492	0,667
196	assessor	0,00465	0,052	95,544	0,667
197	lieutenant	0,00465	0,052	95,596	0,667
198	marshal	0,00465	0,052	95,648	0,667
199	Member of Parliament	0,00465	0,052	95,699	0,667
200	servant	0,00465	0,052	95,751	0,667
201	showman	0,00465	0,052	95,803	0,667
202	technician	0,00465	0,052	95,855	0,667
203	pharmacist	0,00440	0,104	95,959	1,333
204	bookkeeper	0,00433	0,052	96,010	0,667
205	construction worker	0,00433	0,052	96,062	0,667
206	King	0,00433	0,052	96,114	0,667
207	model	0,00433	0,052	96,166	0,667
208	novelist	0,00433	0,052	96,218	0,667
209	policewoman	0,00433	0,052	96,269	0,667
210	social worker	0,00433	0,052	96,321	0,667
211	cameraman	0,00403	0,052	96,373	0,667
212	computer scientist	0,00403	0,052	96,425	0,667
213	factory worker	0,00403	0,052	96,477	0,667
214	gynaecologist	0,00403	0,052	96,528	0,667
215	househusband	0,00403	0,052	96,580	0,667
216	job holder	0,00403	0,052	96,632	0,667
217	librarian	0,00403	0,052	96,684	0,667
218	marketing manager	0,00403	0,052	96,736	0,667
219	master	0,00403	0,052	96,788	0,667
220	player	0,00403	0,052	96,839	0,667
221	psychiatrist	0,00403	0,052	96,891	0,667
222	salesperson	0,00403	0,052	96,943	0,667
223	sculptor	0,00403	0,052	96,995	0,667
224	community manager	0,00375	0,052	97,047	0,667
225	consultant	0,00375	0,052	97,098	0,667
226	ear nose and throat specialist	0,00375	0,052	97,150	0,667
227	fat cat	0,00375	0,052	97,202	0,667
228	film-maker	0,00375	0,052	97,254	0,667
229	oculist	0,00375	0,052	97,306	0,667
230	pensioner	0,00375	0,052	97,358	0,667
231	soloist	0,00375	0,052	97,409	0,667
232	witch doctor	0,00375	0,052	97,461	0,667
233	thief	0,00371	0,104	97,565	1,333
234	counsellor	0,00349	0,052	97,617	0,667
235	greengrocer	0,00349	0,052	97,668	0,667

236	neurologist	0,00349	0,052	97,720	0,667
237	official	0,00349	0,052	97,772	0,667
238	prime minister	0,00349	0,052	97,824	0,667
239	tester	0,00349	0,052	97,876	0,667
240	allergist	0,00325	0,052	97,927	0,667
241	blue-collar worker	0,00325	0,052	97,979	0,667
242	decorator	0,00325	0,052	98,031	0,667
243	fashion model	0,00325	0,052	98,083	0,667
244	fighter pilot	0,00325	0,052	98,135	0,667
245	grocer	0,00325	0,052	98,187	0,667
246	administrator	0,00302	0,052	98,238	0,667
247	mother	0,00302	0,052	98,290	0,667
248	paediatrician	0,00302	0,052	98,342	0,667
249	white-collar worker	0,00302	0,052	98,394	0,667
250	hacker	0,00281	0,052	98,446	0,667
251	principal	0,00281	0,052	98,497	0,667
252	public servant	0,00281	0,052	98,549	0,667
253	administrative assistant	0,00262	0,052	98,601	0,667
254	adviser	0,00262	0,052	98,653	0,667
255	costume designer	0,00262	0,052	98,705	0,667
256	environmentalist	0,00262	0,052	98,756	0,667
257	film star	0,00262	0,052	98,808	0,667
258	troubadour	0,00262	0,052	98,860	0,667
259	anchorman	0,00244	0,052	98,912	0,667
260	biologist	0,00244	0,052	98,964	0,667
261	buyer	0,00244	0,052	99,016	0,667
262	domestic worker	0,00244	0,052	99,067	0,667
263	kitchen maid	0,00244	0,052	99,119	0,667
264	midwife	0,00244	0,052	99,171	0,667
265	attendant	0,00227	0,052	99,223	0,667
266	broadcaster	0,00227	0,052	99,275	0,667
267	film critic	0,00227	0,052	99,326	0,667
268	lifesaver	0,00227	0,052	99,378	0,667
269	porter	0,00227	0,052	99,430	0,667
270	astronaut	0,00211	0,052	99,482	0,667
271	comedian	0,00211	0,052	99,534	0,667
272	runner	0,00211	0,052	99,585	0,667
273	athlete	0,00196	0,052	99,637	0,667
274	iceman	0,00196	0,052	99,689	0,667
275	shoemaker	0,00196	0,052	99,741	0,667
276	geologist	0,00170	0,052	99,793	0,667
277	monk	0,00170	0,052	99,845	0,667
278	masseur	0,00158	0,052	99,896	0,667
279	magician	0,00137	0,052	99,948	0,667
280	lifeguard	0,00128	0,052	100,000	0,667

2.2.9. Computer and internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	mouse	0,63362	5,373	5,373	80,667
2	screen	0,40740	3,819	9,192	57,333
3	keyboard	0,31772	2,753	11,945	41,333
4	Facebook	0,27651	2,620	14,565	39,333
5	laptop	0,21301	1,954	16,519	29,333
6	web	0,21203	1,732	18,250	26,000
7	email	0,19997	2,176	20,426	32,667
8	hardware	0,19080	2,220	22,647	33,333
9	software	0,18819	2,220	24,867	33,333
10	Google	0,17235	1,687	26,554	25,333
11	download	0,16468	1,909	28,464	28,667
12	chat	0,14459	1,465	29,929	22,000
13	computer	0,14403	1,110	31,039	16,667
14	website	0,14396	1,377	32,416	20,667
15	program(me)	0,13715	1,776	34,192	26,667
16	Tuenti	0,13667	1,421	35,613	21,333
17	Twitter	0,12870	1,199	36,812	18,000
18	pendrive	0,12152	1,288	38,099	19,333
19	Internet	0,11981	1,066	39,165	16,000
20	printer	0,11325	1,243	40,409	18,667
21	Windows	0,11202	1,288	41,696	19,333
22	hard disk	0,10958	1,288	42,984	19,333
23	blog	0,10801	1,066	44,050	16,000
24	web page	0,10590	1,110	45,160	16,667
25	Wi-Fi	0,09847	1,110	46,270	16,667
26	social network	0,09000	0,977	47,247	14,667
27	CD(-ROM)	0,08768	1,066	48,313	16,000
28	PC	0,08129	0,710	49,023	10,667
29	Wikipedia	0,07347	0,888	49,911	13,333
30	World Wide Web	0,07234	0,755	50,666	11,333
31	scanner	0,06829	0,755	51,421	11,333
32	Hotmail	0,06475	0,710	52,131	10,667
33	upload	0,06475	0,844	52,975	12,667
34	game	0,06338	0,666	53,641	10,000
35	monitor	0,06193	0,533	54,174	8,000
36	file	0,05978	0,844	55,018	12,667
37	net	0,05973	0,533	55,551	8,000
38	Youtube	0,05943	0,710	56,261	10,667
39	Messenger	0,05719	0,577	56,838	8,667
40	webcam	0,05496	0,622	57,460	9,333
41	virus	0,05263	0,666	58,126	10,000
42	link	0,04837	0,488	58,615	7,333
43	site	0,04770	0,444	59,059	6,667
44	search	0,04682	0,577	59,636	8,667
45	desktop	0,04615	0,400	60,036	6,000

46	(Microsoft) Word	0,04569	0,488	60,524	7,333
47	floppy disk	0,04542	0,488	61,012	7,333
48	folder	0,04476	0,622	61,634	9,333
49	information	0,04306	0,533	62,167	8,000
50	personal computer	0,04305	0,400	62,567	6,000
51	video	0,04249	0,444	63,011	6,667
52	USB	0,04215	0,533	63,544	8,000
53	wireless	0,04031	0,488	64,032	7,333
54	click	0,04030	0,444	64,476	6,667
55	password	0,03995	0,488	64,964	7,333
56	browser	0,03266	0,355	65,320	5,333
57	music	0,03248	0,355	65,675	5,333
58	Apple	0,03235	0,355	66,030	5,333
59	board	0,03198	0,311	66,341	4,667
60	connection	0,03194	0,400	66,741	6,000
61	wiki	0,03149	0,266	67,007	4,000
62	sabe	0,03018	0,400	67,407	6,000
63	netbook	0,02949	0,311	67,718	4,667
64	video game	0,02933	0,266	67,984	4,000
65	online	0,02859	0,400	68,384	6,000
66	firewall	0,02839	0,355	68,739	5,333
67	surf the Net	0,02832	0,266	69,005	4,000
68	page	0,02805	0,311	69,316	4,667
69	DVD	0,02691	0,444	69,760	6,667
70	modem	0,02628	0,355	70,115	5,333
71	bit	0,02582	0,266	70,382	4,000
72	Microsoft	0,02575	0,311	70,693	4,667
73	Macintosh	0,02509	0,266	70,959	4,000
74	network	0,02490	0,266	71,226	4,000
75	Internet Explorer	0,02490	0,266	71,492	4,000
76	motherboard	0,02470	0,266	71,758	4,000
77	fórum	0,02459	0,311	72,069	4,667
78	RAM	0,02434	0,222	72,291	3,333
79	blogger	0,02337	0,266	72,558	4,000
80	Linux	0,02177	0,222	72,780	3,333
81	disk	0,02170	0,311	73,091	4,667
82	compact disc	0,02071	0,222	73,313	3,333
83	film	0,02058	0,266	73,579	4,000
84	delete	0,02053	0,222	73,801	3,333
85	CPU	0,02016	0,222	74,023	3,333
86	operating system	0,01968	0,266	74,290	4,000
87	Gmail	0,01963	0,222	74,512	3,333
88	(Microsoft) Excel	0,01954	0,178	74,689	2,667
89	router	0,01780	0,222	74,911	3,333
90	document	0,01748	0,266	75,178	4,000
91	Skype	0,01685	0,222	75,400	3,333
92	(Microsoft) Access	0,01673	0,222	75,622	3,333
93	Hot Potatoes	0,01664	0,178	75,799	2,667
94	Yahoo	0,01642	0,222	76,021	3,333

95	DVD player	0,01618	0,178	76,199	2,667
96	RAM disk	0,01583	0,178	76,377	2,667
97	speaker	0,01543	0,133	76,510	2,000
98	home	0,01529	0,133	76,643	2,000
99	Mozilla Firefox	0,01525	0,178	76,821	2,667
100	recorder	0,01464	0,222	77,043	3,333
101	cable	0,01446	0,133	77,176	2,000
102	run	0,01435	0,133	77,309	2,000
103	desk	0,01410	0,133	77,442	2,000
104	web browser	0,01334	0,133	77,575	2,000
105	mail	0,01333	0,089	77,664	1,333
106	tablet PC	0,01318	0,178	77,842	2,667
107	send	0,01309	0,222	78,064	3,333
108	central processing unit	0,01287	0,089	78,153	1,333
109	browse	0,01273	0,133	78,286	2,000
110	cookie	0,01253	0,133	78,419	2,000
111	memory	0,01252	0,133	78,552	2,000
112	screensaver	0,01250	0,133	78,686	2,000
113	database	0,01215	0,178	78,863	2,667
114	copy	0,01214	0,222	79,085	3,333
115	application	0,01211	0,133	79,218	2,000
116	design	0,01198	0,089	79,307	1,333
117	WikiLeaks	0,01198	0,089	79,396	1,333
118	charger	0,01187	0,133	79,529	2,000
119	byte	0,01184	0,133	79,663	2,000
120	wire	0,01183	0,178	79,840	2,667
121	print	0,01179	0,133	79,973	2,000
122	paste	0,01174	0,222	80,195	3,333
123	eMule	0,01168	0,178	80,373	2,667
124	Trojan	0,01164	0,133	80,506	2,000
125	server	0,01161	0,133	80,639	2,000
126	execute	0,01160	0,133	80,773	2,000
127	start	0,01155	0,089	80,861	1,333
128	key	0,01138	0,178	81,039	2,667
129	social media	0,01137	0,133	81,172	2,000
130	communication	0,01132	0,089	81,261	1,333
131	InfoJobs	0,01121	0,089	81,350	1,333
132	window	0,01121	0,089	81,439	1,333
133	photo	0,01090	0,133	81,572	2,000
134	electronic book	0,01077	0,089	81,661	1,333
135	Bluetooth	0,01061	0,133	81,794	2,000
136	system	0,01061	0,133	81,927	2,000
137	dictionary	0,01052	0,133	82,060	2,000
138	news	0,01043	0,089	82,149	1,333
139	hacker	0,01025	0,178	82,327	2,667
140	battery	0,01016	0,133	82,460	2,000
141	graphical display unit	0,01012	0,133	82,593	2,000

142	headphones	0,01008	0,133	82,726	2,000
143	programming language	0,01000	0,089	82,815	1,333
144	Java	0,00965	0,089	82,904	1,333
145	post	0,00947	0,133	83,037	2,000
146	Wordreference	0,00935	0,133	83,171	2,000
147	Paint	0,00929	0,089	83,259	1,333
148	open	0,00903	0,089	83,348	1,333
149	Series Yonkis	0,00903	0,089	83,437	1,333
150	joystick	0,00899	0,089	83,526	1,333
151	play	0,00886	0,089	83,615	1,333
152	format	0,00885	0,133	83,748	2,000
153	lock	0,00880	0,089	83,837	1,333
154	(Microsoft) Office	0,00875	0,089	83,925	1,333
155	(Microsoft) PowerPoint	0,00869	0,089	84,014	1,333
156	(Adobe) Photoshop	0,00866	0,089	84,103	1,333
157	Moodle	0,00862	0,089	84,192	1,333
158	sound card	0,00849	0,089	84,281	1,333
159	tool	0,00825	0,133	84,414	2,000
160	WebQuest	0,00819	0,089	84,503	1,333
161	keypad	0,00808	0,089	84,591	1,333
162	megabyte	0,00808	0,089	84,680	1,333
163	Universal Serial Bus	0,00802	0,089	84,769	1,333
164	Memory Stick	0,00795	0,133	84,902	2,000
165	image	0,00782	0,089	84,991	1,333
166	login	0,00781	0,089	85,080	1,333
167	pornography	0,00781	0,089	85,169	1,333
168	ADSL	0,00758	0,089	85,258	1,333
169	book	0,00752	0,089	85,346	1,333
170	El País	0,00736	0,089	85,435	1,333
171	install	0,00736	0,089	85,524	1,333
172	plug	0,00736	0,089	85,613	1,333
173	memory card	0,00732	0,133	85,746	2,000
174	Blue-ray Disc	0,00724	0,089	85,835	1,333
175	eBay	0,00692	0,089	85,924	1,333
176	diskette	0,00684	0,089	86,012	1,333
177	enter	0,00678	0,089	86,101	1,333
178	work	0,00678	0,089	86,190	1,333
179	podcast	0,00677	0,089	86,279	1,333
180	develop	0,00667	0,044	86,323	0,667
181	networking	0,00667	0,044	86,368	0,667
182	pop-up menu	0,00667	0,044	86,412	0,667
183	porn page	0,00667	0,044	86,456	0,667
184	typing	0,00667	0,044	86,501	0,667
185	newspaper	0,00665	0,089	86,590	1,333
186	text processor	0,00665	0,089	86,679	1,333
187	research	0,00662	0,089	86,767	1,333
188	spam	0,00651	0,089	86,856	1,333

189	shopping	0,00648	0,089	86,945	1,333
190	broadband	0,00637	0,089	87,034	1,333
191	hub	0,00627	0,089	87,123	1,333
192	Myspace	0,00627	0,089	87,211	1,333
193	cassette	0,00620	0,044	87,256	0,667
194	globalization	0,00620	0,044	87,300	0,667
195	Google Reader	0,00620	0,044	87,345	0,667
196	Grooveshark	0,00620	0,044	87,389	0,667
197	landscape format	0,00620	0,044	87,433	0,667
198	mass media	0,00620	0,044	87,478	0,667
199	programming	0,00620	0,044	87,522	0,667
200	PSP	0,00620	0,044	87,567	0,667
201	log in	0,00603	0,089	87,655	1,333
202	edit	0,00593	0,089	87,744	1,333
203	Bill Gates	0,00584	0,089	87,833	1,333
204	banner	0,00577	0,044	87,877	0,667
205	desktop computer	0,00577	0,044	87,922	0,667
206	development	0,00577	0,044	87,966	0,667
207	display	0,00577	0,044	88,011	0,667
208	flash memory	0,00577	0,044	88,055	0,667
209	information technology	0,00577	0,044	88,099	0,667
210	picture	0,00577	0,044	88,144	0,667
211	gigabyte	0,00576	0,089	88,233	1,333
212	data	0,00564	0,089	88,321	1,333
213	advertisement	0,00563	0,089	88,410	1,333
214	resource	0,00543	0,089	88,499	1,333
215	backup copy	0,00537	0,044	88,544	0,667
216	downloading	0,00537	0,044	88,588	0,667
217	instruction	0,00537	0,044	88,632	0,667
218	MSN	0,00537	0,044	88,677	0,667
219	option	0,00537	0,044	88,721	0,667
220	piracy	0,00537	0,044	88,766	0,667
221	terabyte	0,00536	0,089	88,854	1,333
222	keyword	0,00529	0,089	88,943	1,333
223	erase	0,00508	0,089	89,032	1,333
224	applet	0,00500	0,044	89,076	0,667
225	code	0,00500	0,044	89,121	0,667
226	error	0,00500	0,044	89,165	0,667
227	Google Translate	0,00500	0,044	89,210	0,667
228	HTML	0,00500	0,044	89,254	0,667
229	IP	0,00500	0,044	89,298	0,667
230	label	0,00500	0,044	89,343	0,667
231	network host	0,00500	0,044	89,387	0,667
232	school	0,00500	0,044	89,432	0,667
233	slide	0,00500	0,044	89,476	0,667
234	table board	0,00500	0,044	89,520	0,667
235	user	0,00500	0,044	89,565	0,667
236	attach	0,00496	0,089	89,654	1,333

237	audio	0,00465	0,044	89,698	0,667
238	CD(-ROM) player	0,00465	0,044	89,742	0,667
239	Interworking Function	0,00465	0,044	89,787	0,667
240	Spotify	0,00465	0,044	89,831	0,667
241	toolbar	0,00465	0,044	89,876	0,667
242	unit of energy	0,00465	0,044	89,920	0,667
243	URL	0,00465	0,044	89,964	0,667
244	reading	0,00440	0,089	90,053	1,333
245	Big Brother	0,00433	0,044	90,098	0,667
246	bookmark	0,00433	0,044	90,142	0,667
247	component	0,00433	0,044	90,187	0,667
248	computer game	0,00433	0,044	90,231	0,667
249	dot	0,00433	0,044	90,275	0,667
250	drive	0,00433	0,044	90,320	0,667
251	fax	0,00433	0,044	90,364	0,667
252	Groupon	0,00433	0,044	90,409	0,667
253	Guadalinux	0,00433	0,044	90,453	0,667
254	ink	0,00433	0,044	90,497	0,667
255	ISP	0,00433	0,044	90,542	0,667
256	mainframe computer	0,00433	0,044	90,586	0,667
257	profile	0,00433	0,044	90,631	0,667
258	signal	0,00433	0,044	90,675	0,667
259	switch on	0,00420	0,089	90,764	1,333
260	switch off	0,00414	0,089	90,853	1,333
261	Acer	0,00403	0,044	90,897	0,667
262	cinema	0,00403	0,044	90,941	0,667
263	compile	0,00403	0,044	90,986	0,667
264	computer security	0,00403	0,044	91,030	0,667
265	El Mundo	0,00403	0,044	91,075	0,667
266	IT assistant	0,00403	0,044	91,119	0,667
267	look for	0,00403	0,044	91,163	0,667
268	mouse pad	0,00403	0,044	91,208	0,667
269	My PC	0,00403	0,044	91,252	0,667
270	notebook	0,00403	0,044	91,297	0,667
271	press	0,00403	0,044	91,341	0,667
272	screener	0,00403	0,044	91,385	0,667
273	spyware	0,00403	0,044	91,430	0,667
274	The Sims	0,00403	0,044	91,474	0,667
275	The Times	0,00403	0,044	91,519	0,667
276	thread	0,00403	0,044	91,563	0,667
277	video card	0,00403	0,044	91,607	0,667
278	virus removal tool	0,00403	0,044	91,652	0,667
279	solution	0,00390	0,089	91,741	1,333
280	number	0,00379	0,089	91,829	1,333
281	ASUS	0,00375	0,044	91,874	0,667

282	button	0,00375	0,044	91,918	0,667
283	CD(-ROM) unit	0,00375	0,044	91,963	0,667
284	Cinetube	0,00375	0,044	92,007	0,667
285	close	0,00375	0,044	92,052	0,667
286	eXeLearning	0,00375	0,044	92,096	0,667
287	high definition	0,00375	0,044	92,140	0,667
288	huge	0,00375	0,044	92,185	0,667
289	insert	0,00375	0,044	92,229	0,667
290	Internet Protocol	0,00375	0,044	92,274	0,667
291	malware	0,00375	0,044	92,318	0,667
292	office suite	0,00375	0,044	92,362	0,667
293	parental control software	0,00375	0,044	92,407	0,667
294	plug-in	0,00375	0,044	92,451	0,667
295	portfolio	0,00375	0,044	92,496	0,667
296	recorderer	0,00375	0,044	92,540	0,667
297	to google	0,00375	0,044	92,584	0,667
298	web application	0,00375	0,044	92,629	0,667
299	source	0,00374	0,089	92,718	1,333
300	add	0,00366	0,089	92,806	1,333
301	load	0,00355	0,089	92,895	1,333
302	fix	0,00349	0,044	92,940	0,667
303	hypertext	0,00349	0,044	92,984	0,667
304	instant messaging	0,00349	0,044	93,028	0,667
305	message	0,00349	0,044	93,073	0,667
306	open source	0,00349	0,044	93,117	0,667
307	pone	0,00349	0,044	93,162	0,667
308	printed circuit board	0,00349	0,044	93,206	0,667
309	reset	0,00349	0,044	93,250	0,667
310	shift	0,00349	0,044	93,295	0,667
311	video call	0,00349	0,044	93,339	0,667
312	widescreen	0,00349	0,044	93,384	0,667
313	address	0,00325	0,044	93,428	0,667
314	article	0,00325	0,044	93,472	0,667
315	band	0,00325	0,044	93,517	0,667
316	carpet	0,00325	0,044	93,561	0,667
317	computation	0,00325	0,044	93,606	0,667
318	free time	0,00325	0,044	93,650	0,667
319	gate	0,00325	0,044	93,694	0,667
320	interface	0,00325	0,044	93,739	0,667
321	iPad	0,00325	0,044	93,783	0,667
322	photography	0,00325	0,044	93,828	0,667
323	return	0,00325	0,044	93,872	0,667
324	search process	0,00325	0,044	93,917	0,667
325	SPSS	0,00325	0,044	93,961	0,667
326	tab key	0,00325	0,044	94,005	0,667

327	technology	0,00325	0,044	94,050	0,667
328	touchpad	0,00325	0,044	94,094	0,667
329	typewriter	0,00325	0,044	94,139	0,667
330	wireless mouse	0,00325	0,044	94,183	0,667
331	zip	0,00325	0,044	94,227	0,667
332	chat room	0,00302	0,044	94,272	0,667
333	control	0,00302	0,044	94,316	0,667
334	directory	0,00302	0,044	94,361	0,667
335	documentary	0,00302	0,044	94,405	0,667
336	Dropbox	0,00302	0,044	94,449	0,667
337	fill	0,00302	0,044	94,494	0,667
338	friend	0,00302	0,044	94,538	0,667
339	holiday	0,00302	0,044	94,583	0,667
340	HP	0,00302	0,044	94,627	0,667
341	innovation	0,00302	0,044	94,671	0,667
342	linkage	0,00302	0,044	94,716	0,667
343	online banking	0,00302	0,044	94,760	0,667
344	programmer	0,00302	0,044	94,805	0,667
345	received	0,00302	0,044	94,849	0,667
346	SlideShare	0,00302	0,044	94,893	0,667
347	stop	0,00302	0,044	94,938	0,667
348	trackpad	0,00302	0,044	94,982	0,667
349	translator	0,00302	0,044	95,027	0,667
350	unzip	0,00302	0,044	95,071	0,667
351	velocity	0,00302	0,044	95,115	0,667
352	view	0,00302	0,044	95,160	0,667
353	worm	0,00302	0,044	95,204	0,667
354	Adobe	0,00281	0,044	95,249	0,667
355	box	0,00281	0,044	95,293	0,667
356	camera	0,00281	0,044	95,337	0,667
357	dynamic random access memory	0,00281	0,044	95,382	0,667
358	electronic data processing	0,00281	0,044	95,426	0,667
359	escape	0,00281	0,044	95,471	0,667
360	feed	0,00281	0,044	95,515	0,667
361	Google Chrome	0,00281	0,044	95,560	0,667
362	installation	0,00281	0,044	95,604	0,667
363	microphone	0,00281	0,044	95,648	0,667
364	new technologies	0,00281	0,044	95,693	0,667
365	peripheral	0,00281	0,044	95,737	0,667
366	protocol	0,00281	0,044	95,782	0,667
367	review	0,00281	0,044	95,826	0,667
368	television	0,00281	0,044	95,870	0,667
369	virtual	0,00281	0,044	95,915	0,667
370	configuration	0,00262	0,044	95,959	0,667

371	ICT	0,00262	0,044	96,004	0,667
372	laser	0,00262	0,044	96,048	0,667
373	loudspeaker	0,00262	0,044	96,092	0,667
374	MS-DOS	0,00262	0,044	96,137	0,667
375	newsletter	0,00262	0,044	96,181	0,667
376	resolution	0,00262	0,044	96,226	0,667
377	sex	0,00262	0,044	96,270	0,667
378	webcast	0,00262	0,044	96,314	0,667
379	MP3	0,00255	0,089	96,403	1,333
380	auricular	0,00244	0,044	96,448	0,667
381	front page	0,00244	0,044	96,492	0,667
382	function	0,00244	0,044	96,536	0,667
383	language	0,00244	0,044	96,581	0,667
384	nickname	0,00244	0,044	96,625	0,667
385	online training	0,00244	0,044	96,670	0,667
386	pixel	0,00244	0,044	96,714	0,667
387	presentation	0,00244	0,044	96,758	0,667
388	rendering	0,00244	0,044	96,803	0,667
389	serial	0,00244	0,044	96,847	0,667
390	share a file	0,00244	0,044	96,892	0,667
391	socket	0,00244	0,044	96,936	0,667
392	spreadsheet	0,00244	0,044	96,980	0,667
393	color	0,00227	0,044	97,025	0,667
394	LCD screen	0,00227	0,044	97,069	0,667
395	LED	0,00227	0,044	97,114	0,667
396	object	0,00227	0,044	97,158	0,667
397	plotter	0,00227	0,044	97,202	0,667
398	post a comment	0,00227	0,044	97,247	0,667
399	tutorial	0,00227	0,044	97,291	0,667
400	write	0,00227	0,044	97,336	0,667
401	electronic mail	0,00211	0,044	97,380	0,667
402	informatics	0,00211	0,044	97,425	0,667
403	job	0,00211	0,044	97,469	0,667
404	letter	0,00211	0,044	97,513	0,667
405	massively multiplayer online game	0,00211	0,044	97,558	0,667
406	series	0,00211	0,044	97,602	0,667
407	social networking website	0,00211	0,044	97,647	0,667
408	sound	0,00211	0,044	97,691	0,667
409	TFT-LCD screen	0,00211	0,044	97,735	0,667
410	USB port	0,00211	0,044	97,780	0,667
411	visual	0,00211	0,044	97,824	0,667
412	chip	0,00196	0,044	97,869	0,667
413	computer programmer	0,00196	0,044	97,913	0,667

414	create	0,00196	0,044	97,957	0,667
415	help	0,00196	0,044	98,002	0,667
416	icon	0,00196	0,044	98,046	0,667
417	JDownloader	0,00196	0,044	98,091	0,667
418	role play	0,00196	0,044	98,135	0,667
419	set up	0,00196	0,044	98,179	0,667
420	share	0,00196	0,044	98,224	0,667
421	subscribe	0,00196	0,044	98,268	0,667
422	under maintenance	0,00196	0,044	98,313	0,667
423	useful	0,00196	0,044	98,357	0,667
424	wall	0,00196	0,044	98,401	0,667
425	computing task	0,00183	0,044	98,446	0,667
426	cut	0,00183	0,044	98,490	0,667
427	entertainment	0,00183	0,044	98,535	0,667
428	Google Maps	0,00183	0,044	98,579	0,667
429	PDF	0,00183	0,044	98,623	0,667
430	reboot	0,00183	0,044	98,668	0,667
431	search engine	0,00183	0,044	98,712	0,667
432	security	0,00183	0,044	98,757	0,667
433	trash	0,00183	0,044	98,801	0,667
434	wireless network	0,00183	0,044	98,845	0,667
435	computer science	0,00170	0,044	98,890	0,667
436	domain	0,00170	0,044	98,934	0,667
437	encryption	0,00170	0,044	98,979	0,667
438	input	0,00170	0,044	99,023	0,667
439	multifunctional	0,00170	0,044	99,067	0,667
440	storage	0,00170	0,044	99,112	0,667
441	tag	0,00170	0,044	99,156	0,667
442	torrent file	0,00170	0,044	99,201	0,667
443	university	0,00170	0,044	99,245	0,667
444	contact	0,00158	0,044	99,290	0,667
445	drag	0,00158	0,044	99,334	0,667
446	output	0,00158	0,044	99,378	0,667
447	speed	0,00158	0,044	99,423	0,667
448	fake	0,00147	0,044	99,467	0,667
449	knowledge	0,00147	0,044	99,512	0,667
450	minimize	0,00147	0,044	99,556	0,667
451	telephone line	0,00147	0,044	99,600	0,667
452	vocabulary	0,00147	0,044	99,645	0,667
453	WordPad	0,00147	0,044	99,689	0,667
454	class	0,00137	0,044	99,734	0,667
455	maximize	0,00137	0,044	99,778	0,667
456	Ares	0,00128	0,044	99,822	0,667
457	log out	0,00128	0,044	99,867	0,667
458	processor	0,00128	0,044	99,911	0,667
459	account	0,00111	0,044	99,956	0,667

460	record	0,00111	0,044	100,000	0,667
-----	--------	---------	-------	---------	-------

3. Dictionarios en francés

3.1. Listado de palabras en francés por orden alfabético

3.1.1. Parties du corps humain

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	abdominal	0,01952	0,234	0,234	4,762
2	articulation	0,01221	0,234	0,468	4,762
3	avant-bras	0,03998	0,468	0,937	9,524
4	barbe	0,01111	0,234	1,171	4,762
5	biceps	0,01952	0,234	1,405	4,762
6	bite	0,02587	0,234	1,639	4,762
7	bouche	0,62218	4,215	5,855	85,714
8	bras	0,73239	4,684	10,539	95,238
9	ceinture	0,01777	0,234	10,773	4,762
10	cerveau	0,01405	0,234	11,007	4,762
11	chatte	0,01618	0,234	11,241	4,762
12	cheveu	0,40744	3,279	14,520	66,667
13	cheville	0,19692	1,639	16,159	33,333
14	cil	0,11558	0,937	17,096	19,048
15	cœur	0,18772	1,874	18,970	38,095
16	colonne	0,02978	0,234	19,204	4,762
17	cou	0,37803	2,576	21,780	52,381
18	coude	0,21977	1,874	23,653	38,095
19	crâne	0,04052	0,468	24,122	9,524
20	cuisse	0,08651	0,703	24,824	14,286
21	cul	0,17631	1,405	26,230	28,571
22	dent	0,32621	2,810	29,040	57,143
23	derrière	0,01952	0,234	29,274	4,762
24	doigt	0,55433	3,981	33,255	80,952
25	dos	0,27895	2,342	35,597	47,619
26	épaule	0,26545	2,342	37,939	47,619
27	épine dorsale	0,02144	0,234	38,173	4,762
28	estomac	0,13368	1,639	39,813	33,333
29	fesse	0,04579	0,468	40,281	9,524
30	foie	0,00524	0,234	40,515	4,762
31	gencive	0,02144	0,234	40,749	4,762
32	genou	0,26285	2,342	43,091	47,619
33	globule	0,01473	0,234	43,326	4,762
34	gorge	0,02468	0,234	43,560	4,762
35	hanche	0,10385	0,937	44,496	19,048
36	index	0,00965	0,234	44,731	4,762
37	intestin	0,03570	0,468	45,199	9,524
38	jambe	0,61430	4,215	49,415	85,714
39	joue	0,03365	0,468	49,883	9,524
40	langue	0,24020	2,108	51,991	42,857

41	lèvre	0,09813	0,937	52,927	19,048
42	main	0,78660	4,684	57,611	95,238
43	menton	0,07769	0,703	58,314	14,286
44	moustache	0,08097	0,703	59,016	14,286
45	muscle	0,09374	1,171	60,187	23,810
46	nerf	0,00576	0,234	60,422	4,762
47	nez	0,55404	3,747	64,169	76,190
48	nombril	0,02247	0,234	64,403	4,762
49	œil	0,45634	2,810	67,213	57,143
50	ongle	0,32931	2,576	69,789	52,381
51	oreille	0,61923	4,450	74,239	90,476
52	orteil	0,01696	0,234	74,473	4,762
53	os	0,09524	0,937	75,410	19,048
54	paume	0,02468	0,234	75,644	4,762
55	paupière	0,01279	0,234	75,878	4,762
56	peau	0,05840	0,468	76,347	9,524
57	pectoral	0,02247	0,234	76,581	4,762
58	petit doigt	0,00921	0,234	76,815	4,762
59	phalange	0,03271	0,234	77,049	4,762
60	pied	0,68610	4,684	81,733	95,238
61	poignet	0,01952	0,234	81,967	4,762
62	poil	0,09682	0,937	82,904	19,048
63	poitrine	0,07739	0,937	83,841	19,048
64	pouce	0,04283	0,468	84,309	9,524
65	poumon	0,14211	1,405	85,714	28,571
66	rein	0,03212	0,468	86,183	9,524
67	ride	0,00604	0,234	86,417	4,762
68	sang	0,04252	0,703	87,119	14,286
69	sein	0,02144	0,234	87,354	4,762
70	sourcil	0,09646	0,937	88,290	19,048
71	talon	0,01862	0,234	88,525	4,762
72	tête	0,80564	4,215	92,740	85,714
73	thorax	0,09873	0,937	93,677	19,048
74	tripes	0,01696	0,234	93,911	4,762
75	veine	0,04499	0,468	94,379	9,524
76	ventre	0,28929	2,342	96,721	47,619
77	visage	0,08472	0,468	97,190	9,524
78	yeux	0,44818	2,810	100,000	57,143

3.1.2. Les vêtements

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	anorak	0,08682	0,885	0,885	14,286
2	bague	0,01165	0,295	1,180	4,762
3	ballerine	0,02046	0,295	1,475	4,762
4	bas	0,04335	0,295	1,770	4,762
5	bermuda	0,03271	0,295	2,065	4,762
6	blouse	0,18813	1,475	3,540	23,810
7	blouson	0,16746	1,770	5,310	28,571
8	Bonnet	0,12595	1,180	6,490	19,048
9	botte	0,17674	1,770	8,260	28,571
10	boucle d'oreille	0,01473	0,295	8,555	4,762
11	bracelet	0,03271	0,295	8,850	4,762
12	caleçon	0,13498	1,180	10,029	19,048
13	cap	0,02587	0,295	10,324	4,762
14	cardigan	0,06013	0,590	10,914	9,524
15	casquette	0,13063	1,180	12,094	19,048
16	ceinture	0,06392	0,590	12,684	9,524
17	châle	0,01862	0,295	12,979	4,762
18	chandail	0,05205	0,590	13,569	9,524
19	chapeau	0,34012	2,950	16,519	47,619
20	chaussette	0,48906	4,130	20,649	66,667
21	chaussure	0,61386	5,015	25,664	80,952
22	chaussure à talon				
	haut	0,04335	0,295	25,959	4,762
23	chemise	0,51924	4,425	30,383	71,429
24	chemisette	0,04936	0,590	30,973	9,524
25	chemisier	0,09811	0,885	31,858	14,286
26	claquette	0,01618	0,295	32,153	4,762
27	collant	0,18365	1,770	33,923	28,571
28	collier	0,01544	0,295	34,218	4,762
29	costume	0,04757	0,590	34,808	9,524
30	cravate	0,23423	2,360	37,168	38,095
31	culotte	0,22855	2,065	39,233	33,333
32	décolleté	0,02711	0,295	39,528	4,762
33	demi-bas	0,04136	0,295	39,823	4,762
34	echarpe	0,32215	3,245	43,068	52,381
35	foulard	0,18409	1,770	44,838	28,571
36	frac	0,03766	0,295	45,133	4,762
37	gant	0,34299	3,245	48,378	52,381
38	gilet	0,02468	0,295	48,673	4,762
39	impermeable	0,08190	0,590	49,263	9,524
40	jaquette	0,02468	0,295	49,558	4,762
41	jean	0,38310	3,540	53,097	57,143
42	jersey	0,10643	0,885	53,982	14,286
43	jupe	0,55686	4,130	58,112	66,667
44	laine	0,03121	0,295	58,407	4,762
45	lunettes	0,09645	0,885	59,292	14,286

46	maillot	0,10679	0,885	60,177	14,286
47	maillot de bain	0,02144	0,295	60,472	4,762
48	manche	0,06015	0,590	61,062	9,524
49	manteau	0,38537	3,245	64,307	52,381
50	marcel	0,02841	0,295	64,602	4,762
51	minijupe	0,08497	0,885	65,487	14,286
52	mocassin	0,02144	0,295	65,782	4,762
53	pantalón	0,73391	5,605	71,386	90,476
54	pantalón à pines	0,01952	0,295	71,681	4,762
55	pendentif	0,01221	0,295	71,976	4,762
56	polo	0,02841	0,295	72,271	4,762
57	porte-jarretelles	0,02587	0,295	72,566	4,762
58	pull(-over)	0,33371	2,655	75,221	42,857
59	pyjama	0,13713	1,180	76,401	19,048
60	robe	0,50993	4,130	80,531	66,667
61	sac	0,11022	1,180	81,711	19,048
62	sac à dos	0,04834	0,590	82,301	9,524
63	salopette	0,10727	0,885	83,186	14,286
64	sandale	0,03814	0,590	83,776	9,524
65	serviette	0,02711	0,295	84,071	4,762
66	short	0,08167	0,885	84,956	14,286
67	slip	0,16713	1,475	86,431	23,810
68	smoking	0,03593	0,295	86,726	4,762
69	socquette	0,01777	0,295	87,021	4,762
70	soulier	0,05639	0,590	87,611	9,524
71	soutien-gorge	0,14260	1,475	89,086	23,810
72	string	0,09622	0,885	89,971	14,286
73	survêtement	0,12247	1,180	91,150	19,048
74	sweater	0,02587	0,295	91,445	4,762
75	t(ce)-shirt	0,51507	4,130	95,575	66,667
76	tailleur	0,01696	0,295	95,870	4,762
77	tong	0,01544	0,295	96,165	4,762
78	tricot	0,04335	0,295	96,460	4,762
79	veste	0,41634	3,245	99,705	52,381
80	vêtement	0,04544	0,295	100,000	4,762

3.1.3. La nourriture et les boissons

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	abricot	0,01664	0,230	0,230	4,762
2	alcohol	0,01122	0,230	0,461	4,762
3	aliment	0,03428	0,230	0,691	4,762
4	anchois	0,00921	0,230	0,922	4,762
5	baguette	0,07990	0,691	1,613	14,286
6	banane	0,04274	0,461	2,074	9,524
7	bélier	0,03210	0,230	2,304	4,762
8	beurre	0,14150	0,922	3,226	19,048
9	bière	0,20238	1,382	4,608	28,571
10	bifteck	0,03006	0,230	4,839	4,762
11	biscuit	0,02266	0,461	5,300	9,524
12	boeuf	0,13957	1,152	6,452	23,810
13	bonbon	0,02027	0,230	6,682	4,762
14	boulangerie	0,03006	0,230	6,912	4,762
15	brioche	0,04783	0,461	7,373	9,524
16	cacao	0,02027	0,230	7,604	4,762
17	café	0,16683	1,152	8,756	23,810
18	café au lait	0,01664	0,230	8,986	4,762
19	canard	0,08672	0,461	9,447	9,524
20	canard à l'orange	0,04459	0,230	9,677	4,762
21	cappuccino	0,03661	0,230	9,908	4,762
22	caramel	0,01595	0,461	10,369	9,524
23	carotte	0,10952	0,922	11,290	19,048
24	carpaccio	0,01777	0,230	11,521	4,762
25	cassis	0,01558	0,230	11,751	4,762
26	cava	0,03910	0,230	11,982	4,762
27	céréale	0,09249	0,691	12,673	14,286
28	cerise	0,04493	0,691	13,364	14,286
29	champagne	0,10533	0,691	14,055	14,286
30	champignon	0,01459	0,230	14,286	4,762
31	charcuterie	0,01664	0,230	14,516	4,762
32	chateaubriand	0,01279	0,230	14,747	4,762
33	chewing-gum	0,00808	0,230	14,977	4,762
34	chocolat	0,38887	2,765	17,742	57,143
35	choucroute	0,04762	0,230	17,972	4,762
36	citron	0,07133	0,922	18,894	19,048
37	citronnade	0,04762	0,230	19,124	4,762
38	Coca-Cola	0,17074	1,382	20,507	28,571
39	colin	0,00984	0,230	20,737	4,762
40	confiture	0,15203	1,152	21,889	23,810
41	côte de porc	0,02488	0,461	22,350	9,524
42	côtelette	0,04002	0,461	22,811	9,524
43	courgette	0,01459	0,230	23,041	4,762
44	couscous	0,03428	0,230	23,272	4,762
45	crème fraîche	0,01558	0,230	23,502	4,762

46	crêpe	0,18845	1,843	25,346	38,095
47	crevette	0,01898	0,230	25,576	4,762
48	croissant	0,21457	1,843	27,419	38,095
49	curry	0,01122	0,230	27,650	4,762
50	daurade	0,01122	0,230	27,880	4,762
51	déjeuner	0,04459	0,230	28,111	4,762
52	dessert	0,08879	0,691	28,802	14,286
53	dinde	0,05033	0,461	29,263	9,524
54	dîner	0,04762	0,230	29,493	4,762
55	eau	0,48593	3,687	33,180	76,190
56	eau gazeuse	0,02815	0,230	33,410	4,762
57	eau minérale	0,03006	0,230	33,641	4,762
58	endive au jambon	0,01198	0,230	33,871	4,762
59	entrecôte	0,01050	0,230	34,101	4,762
60	entrée	0,02636	0,230	34,332	4,762
61	épice	0,00984	0,230	34,562	4,762
62	espèce	0,03210	0,230	34,793	4,762
63	Évian	0,00921	0,230	35,023	4,762
64	Fanta	0,03210	0,230	35,253	4,762
65	farine	0,01459	0,230	35,484	4,762
66	foie	0,05126	0,461	35,945	9,524
67	foie gras	0,02311	0,230	36,175	4,762
68	fondue	0,01664	0,230	36,406	4,762
69	fraise	0,04484	0,691	37,097	14,286
70	frangipane	0,01459	0,230	37,327	4,762
71	fricadelle	0,01898	0,230	37,558	4,762
72	frite	0,13010	0,922	38,479	19,048
73	fromage	0,33963	2,535	41,014	52,381
74	fruit	0,24855	1,843	42,857	38,095
75	galette	0,05608	0,691	43,548	14,286
76	gâteau	0,08563	1,152	44,700	23,810
77	girolle	0,01279	0,230	44,931	4,762
78	glace	0,06594	0,691	45,622	14,286
79	goûter	0,01198	0,230	45,853	4,762
80	hamburger	0,03210	0,230	46,083	4,762
81	haricot	0,02287	0,461	46,544	9,524
82	hot-dog	0,03910	0,230	46,774	4,762
83	huile	0,03927	0,461	47,235	9,524
84	jambon	0,19849	1,613	48,848	33,333
85	jus	0,13996	1,382	50,230	28,571
86	jus de fruit	0,02164	0,230	50,461	4,762
87	jus d'orange	0,03941	0,461	50,922	9,524
88	lait	0,37679	2,995	53,917	61,905
89	légume	0,12595	0,922	54,839	19,048
90	lentille	0,03910	0,230	55,069	4,762
91	macaron	0,04459	0,230	55,300	4,762
92	marmelade	0,01898	0,230	55,530	4,762
93	Martini	0,00862	0,230	55,760	4,762
94	melon	0,05461	0,691	56,452	14,286
95	menthe à l'eau	0,02164	0,230	56,682	4,762

96	mousse	0,04887	0,461	57,143	9,524
97	mouton	0,03335	0,461	57,604	9,524
98	mûr	0,01664	0,230	57,834	4,762
99	oeuf	0,12839	1,152	58,986	23,810
100	oignon	0,12399	0,922	59,908	19,048
101	olive	0,03210	0,230	60,138	4,762
102	omelette	0,20417	1,152	61,290	23,810
103	orange	0,19235	1,843	63,134	38,095
104	Orangina	0,07470	0,691	63,825	14,286
105	paëlla	0,01777	0,230	64,055	4,762
106	pain	0,31265	2,304	66,359	47,619
107	pain au chocolat	0,05108	0,461	66,820	9,524
108	pastèque	0,01198	0,230	67,051	4,762
109	pastis	0,08786	0,922	67,972	19,048
110	patate	0,04459	0,230	68,203	4,762
111	pâte	0,14846	1,152	69,355	23,810
112	pâté en croûte	0,00984	0,230	69,585	4,762
113	pêche	0,02164	0,230	69,816	4,762
114	Perrier	0,03210	0,230	70,046	4,762
115	petit pois	0,00862	0,230	70,276	4,762
116	pizza	0,08918	0,461	70,737	9,524
117	poire	0,04026	0,461	71,198	9,524
118	poisson	0,34871	2,995	74,194	61,905
119	poivre	0,02585	0,461	74,654	9,524
120	ponme	0,26512	2,074	76,728	42,857
121	pomme de terre	0,19044	1,382	78,111	28,571
122	porc	0,07147	0,691	78,802	14,286
123	poule	0,03910	0,230	79,032	4,762
124	poulet	0,09719	0,922	79,954	19,048
125	poulet rôti	0,01366	0,230	80,184	4,762
126	poulette	0,04175	0,230	80,415	4,762
127	profiterole	0,01366	0,230	80,645	4,762
128	quiche	0,09935	0,691	81,336	14,286
129	raisin	0,01122	0,230	81,567	4,762
130	ratatouille	0,02248	0,461	82,028	9,524
131	repas	0,03210	0,230	82,258	4,762
132	rhum	0,05642	0,461	82,719	9,524
133	riz	0,09464	0,922	83,641	19,048
134	rôti	0,01459	0,230	83,871	4,762
135	salade	0,11436	0,922	84,793	19,048
136	salade verte	0,03686	0,461	85,253	9,524
137	sardine	0,01279	0,230	85,484	4,762
138	saucisse	0,06073	0,461	85,945	9,524
139	saucisson	0,02027	0,230	86,175	4,762
140	saumon	0,01366	0,230	86,406	4,762
141	sel	0,05341	0,691	87,097	14,286
142	sirop	0,04191	0,461	87,558	9,524
143	sole	0,02027	0,230	87,788	4,762
144	sorbet	0,02815	0,230	88,018	4,762

145	soupe	0,04842	0,461	88,479	9,524
146	spaghetti	0,07837	0,461	88,940	9,524
147	steak	0,04366	0,461	89,401	9,524
148	steak haché	0,00984	0,230	89,631	4,762
149	steak tartare	0,01664	0,230	89,862	4,762
150	sucré	0,02119	0,461	90,323	9,524
151	sucrerie	0,00756	0,230	90,553	4,762
152	tarte	0,05840	0,691	91,244	14,286
153	tartiflette	0,04175	0,230	91,475	4,762
154	thé	0,06320	0,461	91,935	9,524
155	thon	0,02182	0,461	92,396	9,524
156	tomate	0,10020	0,922	93,318	19,048
157	truite	0,01050	0,230	93,548	4,762
158	veau	0,02468	0,230	93,779	4,762
159	viande	0,25566	2,304	96,083	47,619
160	vin	0,19166	1,613	97,696	33,333
161	vin blanc	0,08918	0,461	98,157	9,524
162	vin rouge	0,04175	0,230	98,387	4,762
163	vinaigre	0,04623	0,461	98,848	9,524
164	vodka	0,02027	0,230	99,078	4,762
165	whisky	0,02815	0,230	99,309	4,762
166	yaourt	0,09340	0,691	100,000	14,286

3.1.4. L'école (meubles et matériel)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	agenda	0,04682	0,619	0,619	9,524
2	agrafeuse	0,02388	0,310	0,929	4,762
3	anglais	0,02388	0,310	1,238	4,762
4	année scolaire	0,00817	0,310	1,548	4,762
5	balle	0,00952	0,310	1,858	4,762
6	ballon	0,01028	0,310	2,167	4,762
7	banc	0,02579	0,310	2,477	4,762
8	bâton de colle	0,03246	0,310	2,786	4,762
9	bibliothèque	0,03504	0,310	3,096	4,762
10	biologie	0,02388	0,310	3,406	4,762
11	blanc	0,04718	0,929	4,334	14,286
12	boisson fraîche	0,02049	0,310	4,644	4,762
13	bureau	0,09172	0,619	5,263	9,524
14	cahier	0,37856	4,644	9,907	71,429
15	cahier à grands carreaux	0,02212	0,310	10,217	4,762
16	cahier à petits carreaux	0,02049	0,310	10,526	4,762
17	calculette	0,02784	0,310	10,836	4,762
18	CD	0,01758	0,310	11,146	4,762
19	CDI	0,04410	0,310	11,455	4,762
20	chaise	0,66183	5,882	17,337	90,476
21	chimie	0,01508	0,310	17,647	4,762
22	ciseaux	0,07634	0,929	18,576	14,286
23	clase	0,03785	0,619	19,195	9,524
24	classeur	0,09140	0,929	20,124	14,286
25	clé	0,01198	0,310	20,433	4,762
26	colle	0,04601	0,619	21,053	9,524
27	collège	0,04085	0,310	21,362	4,762
28	compas	0,05790	0,619	21,981	9,524
29	concierge	0,04926	0,619	22,601	9,524
30	corbeille à papier	0,01758	0,310	22,910	4,762
31	couleur	0,04085	0,310	23,220	4,762
32	couloir	0,06473	0,619	23,839	9,524
33	cour	0,03246	0,310	24,149	4,762
34	cour de récréation	0,01628	0,310	24,458	4,762
35	craie	0,08554	1,238	25,697	19,048
36	crayón	0,57072	4,644	30,341	71,429
37	crayon de couleur	0,04874	0,619	30,960	9,524
38	devoir	0,01293	0,310	31,269	4,762
39	dictionnaire	0,01293	0,310	31,579	4,762
40	dossier	0,11456	1,238	32,817	19,048
41	DVD	0,03504	0,310	33,127	4,762
42	école	0,05363	0,619	33,746	9,524
43	école primaire	0,04762	0,310	34,056	4,762

44	effaceur	0,11189	1,548	35,604	23,810
45	élève	0,10349	1,238	36,842	19,048
46	encre	0,04085	0,310	37,152	4,762
47	encre indélébile	0,01758	0,310	37,461	4,762
48	enseignement secondaire	0,04410	0,310	37,771	4,762
49	enseigner	0,00756	0,310	38,080	4,762
50	équerre	0,06451	0,929	39,009	14,286
51	estrade	0,02226	0,619	39,628	9,524
52	étudier	0,00700	0,310	39,938	4,762
53	faculté	0,03504	0,310	40,248	4,762
54	fauteuil	0,07331	0,619	40,867	9,524
55	fenêtre	0,17860	1,858	42,724	28,571
56	feuille	0,11180	1,548	44,272	23,810
57	feutre	0,15970	1,548	45,820	23,810
58	géographie	0,01758	0,310	46,130	4,762
59	gomme	0,16833	2,167	48,297	33,333
60	gymnase	0,04644	0,929	49,226	14,286
61	Internet	0,01758	0,310	49,536	4,762
62	laboratoire	0,01397	0,310	49,845	4,762
63	langue	0,02049	0,310	50,155	4,762
64	livre	0,42230	4,954	55,108	76,190
65	lumière	0,01110	0,310	55,418	4,762
66	lunettes	0,01898	0,310	55,728	4,762
67	lycée	0,06330	0,929	56,656	14,286
68	mathématiques	0,07575	0,929	57,585	14,286
69	matière	0,02388	0,310	57,895	4,762
70	meuble	0,01628	0,310	58,204	4,762
71	mur	0,02579	0,310	58,514	4,762
72	ordinateur	0,18520	2,477	60,991	38,095
73	papier	0,12534	1,238	62,229	19,048
74	physique	0,02579	0,310	62,539	4,762
75	plafond	0,01397	0,310	62,848	4,762
76	plancher	0,02587	0,619	63,467	9,524
77	plume	0,04085	0,310	63,777	4,762
78	plumier	0,00952	0,310	64,087	4,762
79	portable	0,03947	0,619	64,706	9,524
80	porte	0,12358	1,238	65,944	19,048
81	Post-it	0,03006	0,310	66,254	4,762
82	poubelle	0,07358	1,238	67,492	19,048
83	professeur	0,19834	1,858	69,350	28,571
84	projecteur	0,03655	0,619	69,969	9,524
85	pupitre	0,04410	0,310	70,279	4,762
86	radiocassette	0,03246	0,310	70,588	4,762
87	rapporteur	0,02784	0,310	70,898	4,762
88	récréation	0,03783	0,310	71,207	4,762
89	règle	0,15666	1,858	73,065	28,571
90	rétroprojecteur	0,02212	0,310	73,375	4,762
91	ruban de scotch	0,02784	0,310	73,684	4,762
92	sac	0,04762	0,310	73,994	4,762

93	sac à dos	0,10655	1,548	75,542	23,810
94	salle de professeurs	0,04762	0,310	75,851	4,762
95	science	0,01898	0,310	76,161	4,762
96	souris	0,05055	0,619	76,780	9,524
97	stylo	0,59212	5,263	82,043	80,952
98	stylo-bille	0,01898	0,310	82,353	4,762
99	stylo-feutre	0,01758	0,310	82,663	4,762
100	surveillant	0,00557	0,310	82,972	4,762
101	table	0,72141	6,192	89,164	95,238
102	tableau	0,58657	4,954	94,118	76,190
103	tableau digital	0,02212	0,310	94,427	4,762
104	tableau noir	0,01110	0,310	94,737	4,762
105	taille-crayon	0,08657	0,929	95,666	14,286
106	télévision	0,10813	0,929	96,594	14,286
107	TICE	0,01898	0,310	96,904	4,762
108	Tipp-Ex	0,04444	0,619	97,523	9,524
109	toilette	0,03526	0,619	98,142	9,524
110	toit	0,01508	0,310	98,452	4,762
111	trousse	0,08213	0,929	99,381	14,286
112	université	0,03246	0,310	99,690	4,762
113	waters	0,01758	0,310	100,000	4,762

3.1.5. La ville

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	adresse	0,01053	0,257	0,257	4,762
2	aéroport	0,06245	0,771	1,028	14,286
3	appartement	0,10259	0,771	1,799	14,286
4	arbre	0,06882	1,028	2,828	19,048
5	arène	0,02521	0,514	3,342	9,524
6	arrêt de bus	0,04814	0,514	3,856	9,524
7	autobús	0,19028	1,542	5,398	28,571
8	autoroute	0,02010	0,257	5,656	4,762
9	avenue	0,24417	1,799	7,455	33,333
10	banc	0,02680	0,257	7,712	4,762
11	banlieue	0,04818	0,514	8,226	9,524
12	banque	0,02160	0,257	8,483	4,762
13	bar	0,09480	1,285	9,769	23,810
14	bâtiment	0,30404	2,057	11,825	38,095
15	bibliothèque	0,00735	0,257	12,082	4,762
16	bistrot	0,07153	0,771	12,853	14,286
17	Bois (de Boulogne)	0,01131	0,257	13,111	4,762
18	boîte	0,04062	0,514	13,625	9,524
19	boucherie	0,12872	1,285	14,910	23,810
20	bouchon	0,01871	0,257	15,167	4,762
21	boulangerie	0,22319	2,571	17,738	47,619
22	boulevard	0,06451	0,514	18,252	9,524
23	brasserie	0,01620	0,257	18,509	4,762
24	bureau de tabac	0,03838	0,257	18,766	4,762
25	bus	0,04835	0,514	19,280	9,524
26	café	0,02872	0,514	19,794	9,524
27	camión	0,06445	0,514	20,308	9,524
28	Carrefour	0,10427	1,028	21,337	19,048
29	cathédrale	0,03002	0,514	21,851	9,524
30	centre	0,06165	0,514	22,365	9,524
31	centre urbain	0,04762	0,257	22,622	4,762
32	centre ville	0,07111	0,514	23,136	9,524
33	centre-ville	0,03838	0,257	23,393	4,762
34	champ	0,01620	0,257	23,650	4,762
35	château	0,02010	0,257	23,907	4,762
36	chien	0,02494	0,257	24,165	4,762
37	cimetière	0,04878	0,514	24,679	9,524
38	cinema	0,25159	2,571	27,249	47,619
39	cité	0,01871	0,257	27,506	4,762
40	collège	0,03546	0,514	28,021	9,524
41	cordonnier	0,03086	0,514	28,535	9,524
42	courir	0,00980	0,257	28,792	4,762
43	crêperie	0,04124	0,257	29,049	4,762
44	culture	0,01306	0,257	29,306	4,762
45	département	0,01215	0,257	29,563	4,762

46	discothèque	0,06310	0,771	30,334	14,286
47	école	0,12605	1,542	31,877	28,571
48	édifice	0,07163	0,514	32,391	9,524
49	église	0,18193	2,571	34,961	47,619
50	faubourg	0,03572	0,257	35,219	4,762
51	fête	0,00790	0,257	35,476	4,762
52	feu	0,07441	0,514	35,990	9,524
53	feu rouge	0,09212	1,028	37,018	19,048
54	fleur	0,01871	0,257	37,275	4,762
55	fleuve	0,02680	0,257	37,532	4,762
56	fontaine	0,10618	1,028	38,560	19,048
57	fumée	0,03094	0,257	38,817	4,762
58	garage	0,02680	0,257	39,075	4,762
59	gare	0,11354	1,028	40,103	19,048
60	gens	0,01306	0,257	40,360	4,762
61	hôpital	0,09920	1,028	41,388	19,048
62	hôtel	0,04331	0,514	41,902	9,524
63	hôtel de ville	0,10535	0,771	42,674	14,286
64	immeuble	0,09297	0,771	43,445	14,286
65	jardin	0,12335	1,285	44,730	23,810
66	klaxon	0,02494	0,257	44,987	4,762
67	lampadaire	0,00912	0,257	45,244	4,762
68	lycée	0,04667	0,771	46,015	14,286
69	Lyon	0,03838	0,257	46,272	4,762
70	magasin	0,21559	2,057	48,329	38,095
71	mairie	0,26474	2,571	50,900	47,619
72	maison	0,29328	1,799	52,699	33,333
73	Malaga	0,04124	0,257	52,956	4,762
74	manger	0,00912	0,257	53,213	4,762
75	marche	0,03324	0,257	53,470	4,762
76	marché	0,06201	0,771	54,242	14,286
77	mer	0,05559	0,514	54,756	9,524
78	métro	0,14976	1,799	56,555	33,333
79	monument	0,04654	0,514	57,069	9,524
80	mosquée	0,03094	0,257	57,326	4,762
81	moto	0,10152	1,028	58,355	19,048
82	musée	0,13150	1,285	59,640	23,810
83	oiseau	0,03838	0,257	59,897	4,762
84	opéra	0,01508	0,257	60,154	4,762
85	palace	0,02160	0,257	60,411	4,762
86	palais	0,02010	0,257	60,668	4,762
87	panneau de signalisation	0,04889	0,514	61,183	9,524
88	parc	0,13937	1,542	62,725	28,571
89	parking	0,02494	0,257	62,982	4,762
90	passage de piéton	0,01620	0,257	63,239	4,762
91	pâtisserie	0,07272	1,285	64,524	23,810
92	pavillon	0,02321	0,257	64,781	4,762
93	pharmacie	0,03274	0,514	65,296	9,524
94	piéton	0,03572	0,257	65,553	4,762

95	place	0,28357	2,057	67,609	38,095
96	plage	0,03094	0,257	67,866	4,762
97	plombier	0,03324	0,257	68,123	4,762
98	poissonnerie	0,05334	0,514	68,638	9,524
99	pôlice	0,17473	1,542	70,180	28,571
100	pollution	0,05039	0,514	70,694	9,524
101	pompier	0,02160	0,257	70,951	4,762
102	pont	0,04124	0,257	71,208	4,762
103	port	0,02160	0,257	71,465	4,762
104	poste	0,09302	0,771	72,237	14,286
105	promenade	0,00848	0,257	72,494	4,762
106	pub	0,04002	0,514	73,008	9,524
107	quartier	0,11321	1,028	74,036	19,048
108	restaurant	0,17009	1,542	75,578	28,571
109	rivière	0,05758	0,514	76,093	9,524
110	rond-point	0,07114	0,771	76,864	14,286
111	route	0,29707	2,314	79,177	42,857
112	rue	0,57446	3,856	83,033	71,429
113	ruelle	0,08892	0,771	83,805	14,286
114	service social	0,02494	0,257	84,062	4,762
115	soleil	0,03094	0,257	84,319	4,762
116	sortie	0,01620	0,257	84,576	4,762
117	station	0,02879	0,257	84,833	4,762
118	stop	0,03094	0,257	85,090	4,762
119	stress	0,02680	0,257	85,347	4,762
120	supermarché	0,12789	1,542	86,889	28,571
121	taxi	0,15581	1,028	87,918	19,048
122	temps	0,03324	0,257	88,175	4,762
123	théâtre	0,11338	1,542	89,717	28,571
124	train	0,14397	1,285	91,003	23,810
125	tramway	0,01741	0,257	91,260	4,762
126	transport	0,01131	0,257	91,517	4,762
127	travail	0,01215	0,257	91,774	4,762
128	travailler	0,01403	0,257	92,031	4,762
129	trottoir	0,05138	0,771	92,802	14,286
130	université	0,01620	0,257	93,059	4,762
131	usine	0,03128	0,514	93,573	9,524
132	vélo	0,19323	2,057	95,630	38,095
133	vitrine	0,03094	0,257	95,887	4,762
134	voiture	0,54326	3,856	99,743	71,429
135	zone verte	0,01741	0,257	100,000	4,762

3.1.6. Moyens de transport

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	à pied	0,15119	2,008	2,008	23,810
2	autobús	0,07768	0,803	2,811	9,524
3	AVE	0,01758	0,402	3,213	4,762
4	avion	0,70241	8,032	11,245	95,238
5	bagnole	0,02388	0,402	11,647	4,762
6	barque	0,04249	1,205	12,851	14,286
7	bateau	0,55337	8,032	20,884	95,238
8	bicyclette	0,02388	0,402	21,285	4,762
9	bus	0,63406	6,426	27,711	76,190
10	caisse	0,02212	0,402	28,112	4,762
11	calèche	0,01028	0,402	28,514	4,762
12	camión	0,10494	1,205	29,719	14,286
13	car	0,03006	0,402	30,120	4,762
14	chameau	0,02388	0,402	30,522	4,762
15	charrette	0,02212	0,402	30,924	4,762
16	chemin de fer	0,02212	0,402	31,325	4,762
17	cheval	0,30665	4,819	36,145	57,143
18	éléphant	0,02212	0,402	36,546	4,762
19	funiculaire	0,02784	0,402	36,948	4,762
20	fusée	0,01397	0,402	37,349	4,762
21	hélicoptère	0,04777	0,803	38,153	9,524
22	jet	0,01508	0,402	38,554	4,762
23	métro	0,57888	6,827	45,382	80,952
24	minibús	0,07261	1,205	46,586	14,286
25	moto	0,29835	4,418	51,004	52,381
26	motocycle	0,03246	0,402	51,406	4,762
27	motocyclette	0,09258	1,205	52,610	14,286
28	ómnibus	0,02784	0,402	53,012	4,762
29	parachute	0,07575	1,205	54,217	14,286
30	parapente	0,01293	0,402	54,618	4,762
31	patin	0,03006	0,402	55,020	4,762
32	patin à roulettes	0,01758	0,402	55,422	4,762
33	patinette	0,02579	0,402	55,823	4,762
34	péniche	0,02793	0,803	56,627	9,524
35	RER	0,13012	1,606	58,233	19,048
36	roller	0,05395	0,803	59,036	9,524
37	scooter	0,06664	0,803	59,839	9,524
38	side-car	0,05055	0,803	60,643	9,524
39	skate-board	0,02212	0,402	61,044	4,762
40	ski	0,01198	0,402	61,446	4,762
41	sousmarin	0,01758	0,402	61,847	4,762
42	tanque	0,01898	0,402	62,249	4,762
43	taxi	0,23879	2,811	65,060	33,333
44	TER	0,11373	1,205	66,265	14,286
45	TGV	0,26937	3,213	69,478	38,095

46	tracteur	0,02212	0,402	69,880	4,762
47	train	0,68864	7,229	77,108	85,714
48	traineau	0,01110	0,402	77,510	4,762
49	tram(way)	0,32880	4,418	81,928	52,381
50	transatlantique	0,01508	0,402	82,329	4,762
51	trottinette	0,06159	1,205	83,534	14,286
52	vélo	0,73653	8,032	91,566	95,238
53	voiture	0,81069	8,434	100,000	100,000

3.1.7. Jeux et loisirs

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	activité à l'air libre	0,02784	0,355	0,355	4,762
2	aguagym	0,01898	0,355	0,709	4,762
3	alcohol	0,02388	0,355	1,064	4,762
4	alpinisme	0,02212	0,355	1,418	4,762
5	balle	0,00756	0,355	1,773	4,762
6	ballon prisonnier	0,06498	1,064	2,837	14,286
7	bande dessinée	0,02388	0,355	3,191	4,762
8	base-ball	0,10058	1,418	4,610	19,048
9	basket(-ball)	0,57890	5,674	10,284	76,190
10	beach-volley	0,01198	0,355	10,638	4,762
11	bibliothèque	0,03246	0,355	10,993	4,762
12	bilboquet	0,03006	0,355	11,348	4,762
13	biscuit	0,02579	0,355	11,702	4,762
14	Blokus	0,02388	0,355	12,057	4,762
15	Buzz!	0,04410	0,355	12,411	4,762
16	cache-cache	0,04601	0,709	13,121	9,524
17	café	0,06252	0,709	13,830	9,524
18	camping	0,02579	0,355	14,184	4,762
19	cap ou pas cap?	0,04762	0,355	14,539	4,762
20	carte	0,07614	1,064	15,603	14,286
21	centre commercial	0,03783	0,355	15,957	4,762
22	changer	0,01628	0,355	16,312	4,762
23	chanter	0,03504	0,355	16,667	4,762
24	chasse	0,04085	0,355	17,021	4,762
25	chat perché	0,04628	0,709	17,730	9,524
26	cinema	0,55068	5,319	23,050	71,429
27	cirque	0,04410	0,355	23,404	4,762
28	Cluedo	0,09172	0,709	24,113	9,524
29	concert	0,02212	0,355	24,468	4,762
30	courir	0,01758	0,355	24,823	4,762
31	course à pied	0,02388	0,355	25,177	4,762
32	cricket	0,00952	0,355	25,532	4,762
33	crier	0,02049	0,355	25,887	4,762
34	dames	0,04085	0,355	26,241	4,762
35	danser	0,11618	1,773	28,014	23,810
36	dimanche	0,02784	0,355	28,369	4,762
37	disco	0,02784	0,355	28,723	4,762
38	discothèque	0,11414	1,418	30,142	19,048
39	échecs	0,03504	0,355	30,496	4,762
40	école	0,01293	0,355	30,851	4,762
41	enfant	0,01508	0,355	31,206	4,762
42	équitation	0,05634	0,709	31,915	9,524
43	Facebook	0,02212	0,355	32,270	4,762
44	faire du shopping	0,02388	0,355	32,624	4,762
45	faire la fête	0,05983	0,709	33,333	9,524

46	fête	0,14277	1,418	34,752	19,048
47	fitness	0,01397	0,355	35,106	4,762
48	football	0,59800	5,674	40,780	76,190
49	golf	0,01293	0,355	41,135	4,762
50	gymnase	0,03006	0,355	41,489	4,762
51	handball	0,02579	0,355	41,844	4,762
52	Internet	0,14607	1,773	43,617	23,810
53	jeu de l'oie	0,10303	1,064	44,681	14,286
54	jeu de mots	0,02579	0,355	45,035	4,762
55	jeu de société	0,03504	0,355	45,390	4,762
56	jeu vidéo	0,25802	2,837	48,227	38,095
57	la tomate	0,01758	0,355	48,582	4,762
58	lecture	0,04410	0,355	48,936	4,762
59	lire	0,05553	0,709	49,645	9,524
60	livre	0,01898	0,355	50,000	4,762
61	loto	0,04085	0,355	50,355	4,762
62	loto sportif	0,03783	0,355	50,709	4,762
63	ludo	0,04085	0,355	51,064	4,762
64	marelle	0,06568	0,709	51,773	9,524
65	Monopoly	0,17422	1,773	53,546	23,810
66	monument	0,02784	0,355	53,901	4,762
67	musée	0,10076	1,418	55,319	19,048
68	musique	0,12769	1,064	56,383	14,286
69	natation	0,17264	2,482	58,865	33,333
70	naviguer sur Internet	0,03006	0,355	59,220	4,762
71	opéra	0,03677	0,709	59,929	9,524
72	padel	0,05459	1,064	60,993	14,286
73	parc	0,04410	0,355	61,348	4,762
74	patinoire	0,02579	0,355	61,702	4,762
75	pêche	0,03783	0,355	62,057	4,762
76	pétanque	0,03783	0,355	62,411	4,762
77	petit	0,01397	0,355	62,766	4,762
78	Pilates	0,01628	0,355	63,121	4,762
79	piscine	0,02784	0,355	63,475	4,762
80	plage	0,08389	1,064	64,539	14,286
81	Play Station	0,06989	0,709	65,248	9,524
82	plongée	0,02579	0,355	65,603	4,762
83	poker	0,12635	1,418	67,021	19,048
84	poursuite	0,03246	0,355	67,376	4,762
85	Pro Evolution Soccer	0,04762	0,355	67,730	4,762
86	promenade	0,07589	0,709	68,440	9,524
87	promener	0,07928	1,064	69,504	14,286
88	pub	0,06568	0,709	70,213	9,524
89	quarté	0,03504	0,355	70,567	4,762
90	radio	0,03504	0,355	70,922	4,762
91	randonnée	0,07412	1,064	71,986	14,286
92	récréation	0,01198	0,355	72,340	4,762
93	restaurant	0,07288	0,709	73,050	9,524

94	roller	0,03783	0,355	73,404	4,762
95	rugby	0,19964	1,773	75,177	23,810
96	samedi	0,03246	0,355	75,532	4,762
97	s'amuser	0,01898	0,355	75,887	4,762
98	saut à l'élastique	0,03007	0,709	76,596	9,524
99	saut à la corde	0,05452	1,064	77,660	14,286
100	sauter	0,02212	0,355	78,014	4,762
101	scrabble	0,03504	0,355	78,369	4,762
102	série télévisée	0,01628	0,355	78,723	4,762
103	ski	0,06012	0,709	79,433	9,524
104	sortir	0,05003	0,709	80,142	9,524
105	sortir avec les amies	0,01628	0,355	80,496	4,762
106	spinning	0,01508	0,355	80,851	4,762
107	sport	0,15193	1,773	82,624	23,810
108	squash	0,04514	0,709	83,333	9,524
109	Super Mario	0,02784	0,355	83,688	4,762
110	tarot	0,05458	0,709	84,397	9,524
111	télévision	0,05996	0,709	85,106	9,524
112	tennis	0,29060	3,901	89,007	52,381
113	théâtre	0,25565	2,482	91,489	33,333
114	toupie	0,05395	0,709	92,199	9,524
115	tourisme	0,03504	0,355	92,553	4,762
116	Trivial	0,06510	0,709	93,262	9,524
117	vélo	0,14125	2,128	95,390	28,571
118	vendredi	0,03006	0,355	95,745	4,762
119	visiter la ville	0,03006	0,355	96,099	4,762
120	volley-ball	0,17657	1,773	97,872	23,810
121	voyage	0,01758	0,355	98,227	4,762
122	voyager	0,01628	0,355	98,582	4,762
123	weekend	0,03504	0,355	98,936	4,762
124	Wii	0,06869	0,709	99,645	9,524
125	yoga	0,01758	0,355	100,000	4,762

3.1.8. Professions et métiers

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	acteur	0,12432	1,672	1,672	23,810
2	actrice	0,01898	0,334	2,007	4,762
3	agriculteur	0,03246	0,334	2,341	4,762
4	archéologue	0,03504	0,334	2,676	4,762
5	architecte	0,16295	1,672	4,348	23,810
6	avocat	0,10434	2,007	6,355	28,571
7	banquier	0,02049	0,334	6,689	4,762
8	barmaid	0,01508	0,334	7,023	4,762
9	barman	0,09965	1,338	8,361	19,048
10	biologiste	0,02801	0,669	9,030	9,524
11	boucher	0,20824	1,672	10,702	23,810
12	boulangier	0,47595	4,348	15,050	61,905
13	broker	0,02784	0,334	15,385	4,762
14	cassier	0,02579	0,334	15,719	4,762
15	caviste	0,02212	0,334	16,054	4,762
16	chanteur	0,02212	0,334	16,388	4,762
17	charcutier	0,08495	0,669	17,057	9,524
18	charpentier	0,06459	0,669	17,726	9,524
19	chauffeur	0,15826	1,672	19,398	23,810
20	chef	0,17698	2,007	21,405	28,571
21	chef de cuisine	0,03006	0,334	21,739	4,762
22	chef de rang	0,02579	0,334	22,074	4,762
23	chef d'entreprise	0,01293	0,334	22,408	4,762
24	chevalier	0,02579	0,334	22,742	4,762
25	chirurgien	0,02784	0,334	23,077	4,762
26	coiffeur	0,06510	0,669	23,746	9,524
27	commercial	0,04085	0,334	24,080	4,762
28	comptable	0,05173	0,669	24,749	9,524
29	concierge	0,05918	0,669	25,418	9,524
30	conducteur	0,04542	0,669	26,087	9,524
31	conseiller	0,01758	0,334	26,421	4,762
32	constructeur	0,03504	0,334	26,756	4,762
33	cordonnier	0,09590	1,003	27,759	14,286
34	cuisinier	0,30796	3,344	31,104	47,619
35	dentiste	0,14670	1,672	32,776	23,810
36	directeur	0,06252	0,669	33,445	9,524
37	docteur	0,18388	2,007	35,452	28,571
38	économiste	0,03783	0,334	35,786	4,762
39	écrivain	0,02049	0,334	36,120	4,762
40	électricien	0,04444	0,669	36,789	9,524
41	employé	0,01397	0,334	37,124	4,762
42	employé de banque	0,03783	0,334	37,458	4,762
43	épiciér	0,02579	0,334	37,793	4,762
44	étudiant	0,06083	0,669	38,462	9,524
45	facteur	0,04850	1,003	39,465	14,286

46	femme de ménage	0,01293	0,334	39,799	4,762
47	fermier	0,01110	0,334	40,134	4,762
48	fleuriste	0,03807	0,669	40,803	9,524
49	fonctionnaire	0,02212	0,334	41,137	4,762
50	footballeur	0,02388	0,334	41,472	4,762
51	fruitier	0,03006	0,334	41,806	4,762
52	garçon	0,08411	1,003	42,809	14,286
53	gardien de la paix	0,03795	0,669	43,478	9,524
54	gendarme	0,07656	0,669	44,147	9,524
55	guichetier	0,01198	0,334	44,482	4,762
56	gynécologue	0,00882	0,334	44,816	4,762
57	hôte de vol	0,02579	0,334	45,151	4,762
58	hôtesse de l'air	0,00952	0,334	45,485	4,762
59	infirmier	0,16721	2,007	47,492	28,571
60	informaticien	0,08495	0,669	48,161	9,524
61	ingénieur	0,10072	1,338	49,498	19,048
62	interprète	0,05474	1,003	50,502	14,286
63	jardinier	0,12093	1,003	51,505	14,286
64	joueur	0,02388	0,334	51,839	4,762
65	journaliste	0,04764	0,669	52,508	9,524
66	juge	0,03386	0,669	53,177	9,524
67	libraire	0,01898	0,334	53,512	4,762
68	maçon	0,01758	0,334	53,846	4,762
69	maître	0,14665	1,338	55,184	19,048
70	marchand	0,01508	0,334	55,518	4,762
71	maroquinier	0,02388	0,334	55,853	4,762
72	médecin	0,29994	3,344	59,197	47,619
73	militaire	0,06623	0,669	59,866	9,524
74	musicien	0,03970	0,669	60,535	9,524
75	navigateur	0,01628	0,334	60,870	4,762
76	nettoyant	0,02784	0,334	61,204	4,762
77	officier	0,02579	0,334	61,538	4,762
78	opticien	0,01293	0,334	61,873	4,762
79	ouvrier	0,03445	0,669	62,542	9,524
80	papetier	0,02579	0,334	62,876	4,762
81	pâtissier	0,15971	1,672	64,548	23,810
82	paysan	0,03504	0,334	64,883	4,762
83	pêcheur	0,09616	1,003	65,886	14,286
84	pédagogue	0,01028	0,334	66,221	4,762
85	pédicure	0,04085	0,334	66,555	4,762
86	peintre	0,02212	0,334	66,890	4,762
87	pharmacien	0,07797	1,003	67,893	14,286
88	pilote	0,02388	0,334	68,227	4,762
89	pilote d'avion	0,01028	0,334	68,562	4,762
90	plombier	0,20543	2,007	70,569	28,571
91	poissonnier	0,06030	0,669	71,237	9,524
92	policier	0,36628	3,344	74,582	47,619
93	politicien	0,03795	0,669	75,251	9,524
94	pompier	0,31421	3,010	78,261	42,857
95	postier	0,02212	0,334	78,595	4,762

96	proctologue	0,03783	0,334	78,930	4,762
97	professeur	0,91321	7,023	85,953	100,000
98	psychologue	0,03504	0,334	86,288	4,762
99	réceptionniste	0,01898	0,334	86,622	4,762
100	récolteur	0,01758	0,334	86,957	4,762
101	retraité	0,02212	0,334	87,291	4,762
102	routier	0,03006	0,334	87,625	4,762
103	sanitaire	0,02784	0,334	87,960	4,762
104	secrétaire	0,05832	0,669	88,629	9,524
105	servante	0,03783	0,334	88,963	4,762
106	serveur	0,14097	2,341	91,304	33,333
107	sommelier	0,05395	0,669	91,973	9,524
108	spécialiste	0,02212	0,334	92,308	4,762
109	styliste	0,03504	0,334	92,642	4,762
110	surveillant	0,01397	0,334	92,977	4,762
111	technicien	0,01508	0,334	93,311	4,762
112	tendeur	0,03006	0,334	93,645	4,762
113	traducteur	0,16170	2,341	95,987	33,333
114	traumatologue	0,00952	0,334	96,321	4,762
115	valet	0,02388	0,334	96,656	4,762
116	vendeur	0,16374	2,341	98,997	33,333
117	vétérinaire	0,02388	0,334	99,331	4,762
118	videur	0,01758	0,334	99,666	4,762
119	voleur	0,02388	0,334	100,000	4,762

3.1.9. Ordinateurs et Internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	(e-)mail	0,09415	1,270	1,270	19,048
2	(Microsoft) PowerPoint	0,04666	0,952	2,222	14,286
3	(Microsoft) Word	0,06184	0,952	3,175	14,286
4	Adobe Photoshop	0,02784	0,317	3,492	4,762
5	adresse	0,01758	0,317	3,810	4,762
6	antivirus	0,04438	0,952	4,762	14,286
7	Ares	0,01898	0,317	5,079	4,762
8	blog	0,02049	0,317	5,397	4,762
9	Busuu	0,02388	0,317	5,714	4,762
10	câble	0,15467	1,587	7,302	23,810
11	campus virtuel	0,01110	0,317	7,619	4,762
12	CD	0,13585	1,270	8,889	19,048
13	chat	0,16909	1,905	10,794	28,571
14	chatter	0,12789	1,587	12,381	23,810
15	CityVille	0,01198	0,317	12,698	4,762
16	clavier	0,60541	5,079	17,778	76,190
17	clé USB	0,18674	1,905	19,683	28,571
18	coller	0,05716	0,635	20,317	9,524
19	computer	0,04410	0,317	20,635	4,762
20	connexion	0,01628	0,317	20,952	4,762
21	copier	0,06172	0,635	21,587	9,524
22	courriel	0,01898	0,317	21,905	4,762
23	courrier électronique	0,09493	1,270	23,175	19,048
24	définition	0,03504	0,317	23,492	4,762
25	dictionnaire	0,00952	0,317	23,810	4,762
26	disque dur	0,01397	0,317	24,127	4,762
27	disque dur externe	0,01293	0,317	24,444	4,762
28	disquette	0,07768	0,635	25,079	9,524
29	domaine	0,01508	0,317	25,397	4,762
30	dossier	0,07092	0,952	26,349	14,286
31	DVD	0,05295	0,635	26,984	9,524
32	écran	0,51024	4,127	31,111	61,905
33	écran LCD	0,03783	0,317	31,429	4,762
34	effacer	0,04085	0,317	31,746	4,762
35	électromagnétique	0,01898	0,317	32,063	4,762
36	eMule	0,02049	0,317	32,381	4,762
37	enregistrer	0,04410	0,317	32,698	4,762
38	environnement	0,01397	0,317	33,016	4,762
39	espace	0,03504	0,317	33,333	4,762
40	étude	0,01397	0,317	33,651	4,762
41	extension	0,00649	0,317	33,968	4,762
42	Facebook	0,29120	3,492	37,460	52,381
43	film	0,01198	0,317	37,778	4,762
44	fórum	0,05794	0,952	38,730	14,286

45	Google	0,19372	2,540	41,270	38,095
46	graveur	0,01508	0,317	41,587	4,762
47	hardware	0,11456	1,270	42,857	19,048
48	hi-fi	0,02212	0,317	43,175	4,762
49	Hot Potatoes	0,12580	1,905	45,079	28,571
50	Hotmail	0,03504	0,317	45,397	4,762
51	image	0,00557	0,317	45,714	4,762
52	imprimante	0,14032	1,587	47,302	23,810
53	imprimeur	0,02388	0,317	47,619	4,762
54	informatique	0,04762	0,317	47,937	4,762
55	installer	0,03006	0,317	48,254	4,762
56	Internet	0,31311	2,540	50,794	38,095
57	Interplex	0,01028	0,317	51,111	4,762
58	JClic	0,13667	1,905	53,016	28,571
59	jeu	0,08087	0,952	53,968	14,286
60	jeu vidéo	0,04085	0,317	54,286	4,762
61	journal en ligne	0,01397	0,317	54,603	4,762
62	Les Sims	0,01758	0,317	54,921	4,762
63	Linux	0,01898	0,317	55,238	4,762
64	logiciel	0,19338	1,905	57,143	28,571
65	matériel informatique	0,05713	0,635	57,778	9,524
66	mémoire	0,03006	0,317	58,095	4,762
67	messaging	0,02579	0,317	58,413	4,762
68	Microsoft	0,02049	0,317	58,730	4,762
69	Microsoft Excel	0,02579	0,317	59,048	4,762
70	Microsoft Office	0,01198	0,317	59,365	4,762
71	Microsoft Works	0,02212	0,317	59,683	4,762
72	moteur de recherche	0,02579	0,317	60,000	4,762
73	MP4	0,03783	0,317	60,317	4,762
74	MSN	0,02388	0,317	60,635	4,762
75	musique	0,01397	0,317	60,952	4,762
76	navigateur	0,02049	0,317	61,270	4,762
77	naviguer	0,02579	0,317	61,587	4,762
78	netbook	0,03783	0,317	61,905	4,762
79	opérateur	0,02212	0,317	62,222	4,762
80	ordinateur	0,17694	1,270	63,492	19,048
81	ordinateur portable	0,02784	0,317	63,810	4,762
82	page	0,10337	0,952	64,762	14,286
83	page web	0,26384	2,222	66,984	33,333
84	parler	0,02212	0,317	67,302	4,762
85	portable	0,08847	0,635	67,937	9,524
86	poste de travail	0,01628	0,317	68,254	4,762
87	poubelle	0,01508	0,317	68,571	4,762
88	problème	0,01898	0,317	68,889	4,762
89	programme	0,07589	0,635	69,524	9,524
90	PSP	0,04085	0,317	69,841	4,762
91	puce	0,03246	0,317	70,159	4,762
92	questionnaire à	0,01628	0,317	70,476	4,762

	choix multiple				
93	recherche	0,01293	0,317	70,794	4,762
94	réseau	0,03246	0,317	71,111	4,762
95	réseau social	0,11968	1,270	72,381	19,048
96	s'amuser	0,00882	0,317	72,698	4,762
97	scanner	0,10656	1,587	74,286	23,810
98	série	0,04601	0,635	74,921	9,524
99	serveur	0,04410	0,317	75,238	4,762
100	site (web)	0,14077	1,270	76,508	19,048
101	Skype	0,01110	0,317	76,825	4,762
102	software	0,08854	0,952	77,778	14,286
103	souris	0,66379	5,079	82,857	76,190
104	spam	0,03191	0,635	83,492	9,524
105	Spotify	0,09721	1,270	84,762	19,048
106	surfer	0,05634	0,635	85,397	9,524
107	système opératif	0,01758	0,317	85,714	4,762
108	technicien	0,01758	0,317	86,032	4,762
109	téléchargement	0,08545	0,635	86,667	9,524
110	télécharger	0,07566	0,952	87,619	14,286
111	touche	0,09325	0,952	88,571	14,286
112	tour	0,03677	0,635	89,206	9,524
113	Tuenti	0,10495	1,270	90,476	19,048
114	Twitter	0,07480	0,952	91,429	14,286
115	USB	0,03246	0,317	91,746	4,762
116	vidéo	0,03279	0,635	92,381	9,524
117	virus	0,09265	1,587	93,968	23,810
118	web	0,07582	0,952	94,921	14,286
119	Web 2.0	0,00952	0,317	95,238	4,762
120	webcam	0,04336	0,635	95,873	9,524
121	Wi-Fi	0,11470	1,270	97,143	19,048
122	Wii	0,02579	0,317	97,460	4,762
123	Wikipedia	0,06785	0,952	98,413	14,286
124	Windows	0,04634	0,635	99,048	9,524
125	WordReference	0,01508	0,317	99,365	4,762
126	Yahoo	0,01898	0,317	99,683	4,762
127	Youtube	0,03006	0,317	100,000	4,762

3.2. Listado de palabras en francés ordenado según el índice de disponibilidad

3.2.1. Parties du corps humain

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	tête	0,80564	4,215	4,215	85,714
2	main	0,78660	4,684	8,899	95,238
3	bras	0,73239	4,684	20,141	95,238
4	pied	0,68610	4,684	35,129	95,238
5	bouche	0,62218	4,215	41,920	85,714
6	oreille	0,61923	4,450	55,972	90,476
7	jambe	0,61430	4,215	30,445	85,714
8	doigt	0,55433	3,981	15,457	80,952
9	nez	0,55404	3,747	48,712	76,190
10	œil	0,45634	2,810	51,522	57,143
11	yeux	0,44818	2,810	68,150	57,143
12	cheveu	0,40744	3,279	76,112	66,667
13	cou	0,37803	2,576	11,475	52,381
14	ongle	0,32931	2,576	37,705	52,381
15	dent	0,32621	2,810	70,960	57,143
16	ventre	0,28929	2,342	22,482	47,619
17	dos	0,27895	2,342	24,824	47,619
18	épaule	0,26545	2,342	82,670	47,619
19	genou	0,26285	2,342	62,529	47,619
20	langue	0,24020	2,108	44,965	42,857
21	coude	0,21977	1,874	72,834	38,095
22	cheville	0,19692	1,639	77,752	33,333
23	cœur	0,18772	1,874	85,714	38,095
24	cul	0,17631	1,405	26,230	28,571
25	poumon	0,14211	1,405	87,119	28,571
26	estomac	0,13368	1,639	57,611	33,333
27	cil	0,11558	0,937	79,391	19,048
28	hanche	0,10385	0,937	63,466	19,048
29	thorax	0,09873	0,937	64,871	19,048
30	lèvre	0,09813	0,937	42,857	19,048
31	poil	0,09682	0,937	91,335	19,048
32	sourcil	0,09646	0,937	80,328	19,048
33	os	0,09524	0,937	89,696	19,048
34	muscle	0,09374	1,171	59,953	23,810
35	cuisse	0,08651	0,703	88,525	14,286
36	visage	0,08472	0,468	78,454	9,524
37	moustache	0,08097	0,703	98,829	14,286
38	mentón	0,07769	0,703	83,372	14,286
39	poitrine	0,07739	0,937	58,548	19,048
40	peau	0,05840	0,468	87,588	9,524
41	fesse	0,04579	0,468	63,934	9,524
42	veine	0,04499	0,468	90,164	9,524

43	pouce	0,04283	0,468	96,253	9,524
44	sang	0,04252	0,703	93,677	14,286
45	crâne	0,04052	0,468	65,340	9,524
46	avant-bras	0,03998	0,468	92,506	9,524
47	intestin	0,03570	0,468	95,785	9,524
48	joue	0,03365	0,468	92,037	9,524
49	phalange	0,03271	0,234	90,398	4,762
50	rein	0,03212	0,468	95,316	9,524
51	colonne	0,02978	0,234	96,956	4,762
52	bite	0,02587	0,234	94,614	4,762
53	paume	0,02468	0,234	96,721	4,762
54	gorge	0,02468	0,234	97,190	4,762
55	nombril	0,02247	0,234	83,607	4,762
56	pectoral	0,02247	0,234	91,569	4,762
57	épine dorsale	0,02144	0,234	94,379	4,762
58	sein	0,02144	0,234	96,487	4,762
59	gencive	0,02144	0,234	97,424	4,762
60	abdominal	0,01952	0,234	58,782	4,762
61	derrière	0,01952	0,234	87,822	4,762
62	poignet	0,01952	0,234	88,759	4,762
63	biceps	0,01952	0,234	92,740	4,762
64	talon	0,01862	0,234	83,841	4,762
65	ceinture	0,01777	0,234	60,187	4,762
66	orteil	0,01696	0,234	77,986	4,762
67	tripes	0,01696	0,234	92,974	4,762
68	chatte	0,01618	0,234	94,848	4,762
69	globule	0,01473	0,234	93,911	4,762
70	cerveau	0,01405	0,234	97,658	4,762
71	paupière	0,01279	0,234	97,892	4,762
72	articulation	0,01221	0,234	94,145	4,762
73	barbe	0,01111	0,234	98,126	4,762
74	index	0,00965	0,234	99,063	4,762
75	petit doigt	0,00921	0,234	99,297	4,762
76	ride	0,00604	0,234	99,532	4,762
77	nerf	0,00576	0,234	99,766	4,762
78	foie	0,00524	0,234	100,000	4,762

3.2.2. Les vêtements

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	pantalón	0,73391	5,605	5,605	90,476
2	chaussure	0,61386	5,015	10,619	80,952
3	jupe	0,55686	4,130	14,749	66,667
4	chemise	0,51924	4,425	19,174	71,429
5	t(ee)-shirt	0,51507	4,130	23,304	66,667
6	robe	0,50993	4,130	27,434	66,667
7	chaussette	0,48906	4,130	31,563	66,667
8	veste	0,41634	3,245	34,808	52,381
9	manteau	0,38537	3,245	38,053	52,381
10	jean	0,38310	3,540	41,593	57,143
11	gant	0,34299	3,245	44,838	52,381
12	chapeau	0,34012	2,950	47,788	47,619
13	pull(-over)	0,33371	2,655	50,442	42,857
14	echarpe	0,32215	3,245	53,687	52,381
15	cravate	0,23423	2,360	56,047	38,095
16	culotte	0,22855	2,065	58,112	33,333
17	blouse	0,18813	1,475	59,587	23,810
18	foulard	0,18409	1,770	61,357	28,571
19	collant	0,18365	1,770	63,127	28,571
20	botte	0,17674	1,770	64,897	28,571
21	blouson	0,16746	1,770	66,667	28,571
22	slip	0,16713	1,475	68,142	23,810
23	soutien-gorge	0,14260	1,475	69,617	23,810
24	pijama	0,13713	1,180	70,796	19,048
25	caleçon	0,13498	1,180	71,976	19,048
26	casquette	0,13063	1,180	73,156	19,048
27	Bonnet	0,12595	1,180	74,336	19,048
28	survêtement	0,12247	1,180	75,516	19,048
29	sac	0,11022	1,180	76,696	19,048
30	salopette	0,10727	0,885	77,581	14,286
31	maillot	0,10679	0,885	78,466	14,286
32	jersey	0,10643	0,885	79,351	14,286
33	chemisier	0,09811	0,885	80,236	14,286
34	lunettes	0,09645	0,885	81,121	14,286
35	string	0,09622	0,885	82,006	14,286
36	anorak	0,08682	0,885	82,891	14,286
37	minijupe	0,08497	0,885	83,776	14,286
38	imperméable	0,08190	0,590	84,366	9,524
39	short	0,08167	0,885	85,251	14,286
40	ceinture	0,06392	0,590	85,841	9,524
41	manche	0,06015	0,590	86,431	9,524
42	cardigan	0,06013	0,590	87,021	9,524
43	soulier	0,05639	0,590	87,611	9,524
44	chandail	0,05205	0,590	88,201	9,524
45	chemisette	0,04936	0,590	88,791	9,524

46	sac à dos	0,04834	0,590	89,381	9,524
47	costume	0,04757	0,590	89,971	9,524
48	vêtement	0,04544	0,295	90,265	4,762
49	tricot	0,04335	0,295	90,560	4,762
50	bas	0,04335	0,295	90,855	4,762
51	chaussure à talon				
	haut	0,04335	0,295	91,150	4,762
52	demi-bas	0,04136	0,295	91,445	4,762
53	sandale	0,03814	0,590	92,035	9,524
54	frac	0,03766	0,295	92,330	4,762
55	smoking	0,03593	0,295	92,625	4,762
56	bermuda	0,03271	0,295	92,920	4,762
57	bracelet	0,03271	0,295	93,215	4,762
58	laine	0,03121	0,295	93,510	4,762
59	polo	0,02841	0,295	93,805	4,762
60	marcel	0,02841	0,295	94,100	4,762
61	décolleté	0,02711	0,295	94,395	4,762
62	serviette	0,02711	0,295	94,690	4,762
63	sweater	0,02587	0,295	94,985	4,762
64	cap	0,02587	0,295	95,280	4,762
65	porte-jarretelles	0,02587	0,295	95,575	4,762
66	jaquette	0,02468	0,295	95,870	4,762
67	gilet	0,02468	0,295	96,165	4,762
68	maillot de bain	0,02144	0,295	96,460	4,762
69	mocassin	0,02144	0,295	96,755	4,762
70	ballerine	0,02046	0,295	97,050	4,762
71	pantalon à pinces	0,01952	0,295	97,345	4,762
72	châle	0,01862	0,295	97,640	4,762
73	socquette	0,01777	0,295	97,935	4,762
74	tailleur	0,01696	0,295	98,230	4,762
75	claquette	0,01618	0,295	98,525	4,762
76	tong	0,01544	0,295	98,820	4,762
77	collier	0,01544	0,295	99,115	4,762
78	boucle d'oreille	0,01473	0,295	99,410	4,762
79	pendentif	0,01221	0,295	99,705	4,762
80	bague	0,01165	0,295	100,000	4,762

3.2.3. La nourriture et les boissons

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	eau	0,48593	3,687	3,687	76,190
2	chocolat	0,38887	2,765	6,452	57,143
3	lait	0,37679	2,995	9,447	61,905
4	poisson	0,34871	2,995	12,442	61,905
5	fromage	0,33963	2,535	14,977	52,381
6	pain	0,31265	2,304	17,281	47,619
7	ponme	0,26512	2,074	19,355	42,857
8	viande	0,25566	2,304	21,659	47,619
9	fruit	0,24855	1,843	23,502	38,095
10	croissant	0,21457	1,843	25,346	38,095
11	omelette	0,20417	1,152	26,498	23,810
12	bière	0,20238	1,382	27,880	28,571
13	jambon	0,19849	1,613	29,493	33,333
14	orange	0,19235	1,843	31,336	38,095
15	vin	0,19166	1,613	32,949	33,333
16	pomme de terre	0,19044	1,382	34,332	28,571
17	crêpe	0,18845	1,843	36,175	38,095
18	Coca-Cola	0,17074	1,382	37,558	28,571
19	café	0,16683	1,152	38,710	23,810
20	confiture	0,15203	1,152	39,862	23,810
21	pâte	0,14846	1,152	41,014	23,810
22	beurre	0,14150	0,922	41,935	19,048
23	jus	0,13996	1,382	43,318	28,571
24	boeuf	0,13957	1,152	44,470	23,810
25	frite	0,13010	0,922	45,392	19,048
26	oeuf	0,12839	1,152	46,544	23,810
27	légume	0,12595	0,922	47,465	19,048
28	oignon	0,12399	0,922	48,387	19,048
29	salade	0,11436	0,922	49,309	19,048
30	carotte	0,10952	0,922	50,230	19,048
31	champagne	0,10533	0,691	50,922	14,286
32	tomate	0,10020	0,922	51,843	19,048
33	quiche	0,09935	0,691	52,535	14,286
34	poulet	0,09719	0,922	53,456	19,048
35	riz	0,09464	0,922	54,378	19,048
36	yaourt	0,09340	0,691	55,069	14,286
37	céréale	0,09249	0,691	55,760	14,286
38	pizza	0,08918	0,461	56,221	9,524
39	vin blanc	0,08918	0,461	56,682	9,524
40	dessert	0,08879	0,691	57,373	14,286
41	pastis	0,08786	0,922	58,295	19,048
42	canard	0,08672	0,461	58,756	9,524
43	gâteau	0,08563	1,152	59,908	23,810
44	baguette	0,07990	0,691	60,599	14,286
45	spaghetti	0,07837	0,461	61,060	9,524

46	Orangina	0,07470	0,691	61,751	14,286
47	porc	0,07147	0,691	62,442	14,286
48	citron	0,07133	0,922	63,364	19,048
49	glace	0,06594	0,691	64,055	14,286
50	thé	0,06320	0,461	64,516	9,524
51	saucisse	0,06073	0,461	64,977	9,524
52	tarte	0,05840	0,691	65,668	14,286
53	rhum	0,05642	0,461	66,129	9,524
54	galette	0,05608	0,691	66,820	14,286
55	melon	0,05461	0,691	67,512	14,286
56	sel	0,05341	0,691	68,203	14,286
57	foie	0,05126	0,461	68,664	9,524
58	pain au chocolat	0,05108	0,461	69,124	9,524
59	dinde	0,05033	0,461	69,585	9,524
60	mousse	0,04887	0,461	70,046	9,524
61	soupe	0,04842	0,461	70,507	9,524
62	brioche	0,04783	0,461	70,968	9,524
63	choucroute	0,04762	0,230	71,198	4,762
64	citronnade	0,04762	0,230	71,429	4,762
65	dîner	0,04762	0,230	71,659	4,762
66	vinaigre	0,04623	0,461	72,120	9,524
67	cerise	0,04493	0,691	72,811	14,286
68	fraise	0,04484	0,691	73,502	14,286
69	canard à l'orange	0,04459	0,230	73,733	4,762
70	déjeuner	0,04459	0,230	73,963	4,762
71	macaron	0,04459	0,230	74,194	4,762
72	patate	0,04459	0,230	74,424	4,762
73	steak	0,04366	0,461	74,885	9,524
74	banane	0,04274	0,461	75,346	9,524
75	sirop	0,04191	0,461	75,806	9,524
76	poulette	0,04175	0,230	76,037	4,762
77	tartiflette	0,04175	0,230	76,267	4,762
78	vin rouge	0,04175	0,230	76,498	4,762
79	poire	0,04026	0,461	76,959	9,524
80	côtelette	0,04002	0,461	77,419	9,524
81	jus d'orange	0,03941	0,461	77,880	9,524
82	huile	0,03927	0,461	78,341	9,524
83	cava	0,03910	0,230	78,571	4,762
84	hot-dog	0,03910	0,230	78,802	4,762
85	lentille	0,03910	0,230	79,032	4,762
86	poule	0,03910	0,230	79,263	4,762
87	salade verte	0,03686	0,461	79,724	9,524
88	cappuccino	0,03661	0,230	79,954	4,762
89	aliment	0,03428	0,230	80,184	4,762
90	couscous	0,03428	0,230	80,415	4,762
91	mouton	0,03335	0,461	80,876	9,524
92	bélier	0,03210	0,230	81,106	4,762
93	espèce	0,03210	0,230	81,336	4,762
94	Fanta	0,03210	0,230	81,567	4,762
95	hamburger	0,03210	0,230	81,797	4,762

96	olive	0,03210	0,230	82,028	4,762
97	Perrier	0,03210	0,230	82,258	4,762
98	repas	0,03210	0,230	82,488	4,762
99	bifteck	0,03006	0,230	82,719	4,762
100	boulangerie	0,03006	0,230	82,949	4,762
101	eau minérale	0,03006	0,230	83,180	4,762
102	eau gazeuse	0,02815	0,230	83,410	4,762
103	sorbet	0,02815	0,230	83,641	4,762
104	whisky	0,02815	0,230	83,871	4,762
105	entrée	0,02636	0,230	84,101	4,762
106	poivre	0,02585	0,461	84,562	9,524
107	côte de porc	0,02488	0,461	85,023	9,524
108	veau	0,02468	0,230	85,253	4,762
109	foie gras	0,02311	0,230	85,484	4,762
110	haricot	0,02287	0,461	85,945	9,524
111	biscuit	0,02266	0,461	86,406	9,524
112	ratatouille	0,02248	0,461	86,866	9,524
113	ton	0,02182	0,461	87,327	9,524
114	jus de fruit	0,02164	0,230	87,558	4,762
115	menthe à l'eau	0,02164	0,230	87,788	4,762
116	pêche	0,02164	0,230	88,018	4,762
117	sucre	0,02119	0,461	88,479	9,524
118	bonbon	0,02027	0,230	88,710	4,762
119	cacao	0,02027	0,230	88,940	4,762
120	saucisson	0,02027	0,230	89,171	4,762
121	sole	0,02027	0,230	89,401	4,762
122	vodka	0,02027	0,230	89,631	4,762
123	crevette	0,01898	0,230	89,862	4,762
124	fricadelle	0,01898	0,230	90,092	4,762
125	marmelade	0,01898	0,230	90,323	4,762
126	carpaccio	0,01777	0,230	90,553	4,762
127	paëlla	0,01777	0,230	90,783	4,762
128	abricot	0,01664	0,230	91,014	4,762
129	café au lait	0,01664	0,230	91,244	4,762
130	charcuterie	0,01664	0,230	91,475	4,762
131	fondue	0,01664	0,230	91,705	4,762
132	mûr	0,01664	0,230	91,935	4,762
133	steak tartare	0,01664	0,230	92,166	4,762
134	caramel	0,01595	0,461	92,627	9,524
135	cassis	0,01558	0,230	92,857	4,762
136	crème fraîche	0,01558	0,230	93,088	4,762
137	champignon	0,01459	0,230	93,318	4,762
138	courgette	0,01459	0,230	93,548	4,762
139	farine	0,01459	0,230	93,779	4,762
140	frangipane	0,01459	0,230	94,009	4,762
141	rôti	0,01459	0,230	94,240	4,762
142	poulet rôti	0,01366	0,230	94,470	4,762
143	profiterole	0,01366	0,230	94,700	4,762
144	saumon	0,01366	0,230	94,931	4,762

145	chateaubriand	0,01279	0,230	95,161	4,762
146	girolle	0,01279	0,230	95,392	4,762
147	sardine	0,01279	0,230	95,622	4,762
148	endive au jambon	0,01198	0,230	95,853	4,762
149	goûter	0,01198	0,230	96,083	4,762
150	pastèque	0,01198	0,230	96,313	4,762
151	alcohol	0,01122	0,230	96,544	4,762
152	curry	0,01122	0,230	96,774	4,762
153	daurade	0,01122	0,230	97,005	4,762
154	raisin	0,01122	0,230	97,235	4,762
155	entrecôte	0,01050	0,230	97,465	4,762
156	truite	0,01050	0,230	97,696	4,762
157	colin	0,00984	0,230	97,926	4,762
158	épice	0,00984	0,230	98,157	4,762
159	pâté en crôte	0,00984	0,230	98,387	4,762
160	steak haché	0,00984	0,230	98,618	4,762
161	anchois	0,00921	0,230	98,848	4,762
162	Évian	0,00921	0,230	99,078	4,762
163	Martini	0,00862	0,230	99,309	4,762
164	petit pois	0,00862	0,230	99,539	4,762
165	chewing-gum	0,00808	0,230	99,770	4,762
166	sucrerie	0,00756	0,230	100,000	4,762

3.2.4. L'école (meubles et matériel)

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	table	0,72141	6,192	6,192	95,238
2	chaise	0,66183	5,882	12,074	90,476
3	stylo	0,59212	5,263	17,337	80,952
4	tableau	0,58657	4,954	22,291	76,190
5	crayón	0,57072	4,644	26,935	71,429
6	livre	0,42230	4,954	31,889	76,190
7	cahier	0,37856	4,644	36,533	71,429
8	professeur	0,19834	1,858	38,390	28,571
9	ordinateur	0,18520	2,477	40,867	38,095
10	fenêtre	0,17860	1,858	42,724	28,571
11	gomme	0,16833	2,167	44,892	33,333
12	feutre	0,15970	1,548	46,440	23,810
13	règle	0,15666	1,858	48,297	28,571
14	papier	0,12534	1,238	49,536	19,048
15	porte	0,12358	1,238	50,774	19,048
16	dossier	0,11456	1,238	52,012	19,048
17	effaceur	0,11189	1,548	53,560	23,810
18	feuille	0,11180	1,548	55,108	23,810
19	télévision	0,10813	0,929	56,037	14,286
20	sac à dos	0,10655	1,548	57,585	23,810
21	élève	0,10349	1,238	58,824	19,048
22	bureau	0,09172	0,619	59,443	9,524
23	classeur	0,09140	0,929	60,372	14,286
24	taille-crayon	0,08657	0,929	61,300	14,286
25	craie	0,08554	1,238	62,539	19,048
26	trousse	0,08213	0,929	63,467	14,286
27	ciseaux	0,07634	0,929	64,396	14,286
28	mathématiques	0,07575	0,929	65,325	14,286
29	poubelle	0,07358	1,238	66,563	19,048
30	fauteuil	0,07331	0,619	67,183	9,524
31	couloir	0,06473	0,619	67,802	9,524
32	équerre	0,06451	0,929	68,731	14,286
33	lycée	0,06330	0,929	69,659	14,286
34	compas	0,05790	0,619	70,279	9,524
35	école	0,05363	0,619	70,898	9,524
36	souris	0,05055	0,619	71,517	9,524
37	concierge	0,04926	0,619	72,136	9,524
38	crayon de couleur	0,04874	0,619	72,755	9,524
39	école primaire	0,04762	0,310	73,065	4,762
40	sac	0,04762	0,310	73,375	4,762
41	salle de professeurs	0,04762	0,310	73,684	4,762
42	blanc	0,04718	0,929	74,613	14,286
43	agenda	0,04682	0,619	75,232	9,524
44	gymnase	0,04644	0,929	76,161	14,286
45	colle	0,04601	0,619	76,780	9,524

46	Tipp-Ex	0,04444	0,619	77,399	9,524
47	CDI	0,04410	0,310	77,709	4,762
48	enseignement secondaire	0,04410	0,310	78,019	4,762
49	pupitre	0,04410	0,310	78,328	4,762
50	collège	0,04085	0,310	78,638	4,762
51	couleur	0,04085	0,310	78,947	4,762
52	encre	0,04085	0,310	79,257	4,762
53	plume	0,04085	0,310	79,567	4,762
54	portable	0,03947	0,619	80,186	9,524
55	clase	0,03785	0,619	80,805	9,524
56	récréation	0,03783	0,310	81,115	4,762
57	projecteur	0,03655	0,619	81,734	9,524
58	toilette	0,03526	0,619	82,353	9,524
59	bibliothèque	0,03504	0,310	82,663	4,762
60	DVD	0,03504	0,310	82,972	4,762
61	faculté	0,03504	0,310	83,282	4,762
62	bâton de colle	0,03246	0,310	83,591	4,762
63	cour	0,03246	0,310	83,901	4,762
64	radiocassette	0,03246	0,310	84,211	4,762
65	université	0,03246	0,310	84,520	4,762
66	Post-it	0,03006	0,310	84,830	4,762
67	calculatrice	0,02784	0,310	85,139	4,762
68	rapporteur	0,02784	0,310	85,449	4,762
69	ruban de scotch	0,02784	0,310	85,759	4,762
70	plancher	0,02587	0,619	86,378	9,524
71	banc	0,02579	0,310	86,687	4,762
72	mur	0,02579	0,310	86,997	4,762
73	physique	0,02579	0,310	87,307	4,762
74	agrafeuse	0,02388	0,310	87,616	4,762
75	anglais	0,02388	0,310	87,926	4,762
76	biologie	0,02388	0,310	88,235	4,762
77	matière	0,02388	0,310	88,545	4,762
78	estrade	0,02226	0,619	89,164	9,524
79	cahier à grands carreaux	0,02212	0,310	89,474	4,762
80	rétroprojecteur	0,02212	0,310	89,783	4,762
81	tableau digital	0,02212	0,310	90,093	4,762
82	boisson fraîche	0,02049	0,310	90,402	4,762
83	cahier à petits carreaux	0,02049	0,310	90,712	4,762
84	langue	0,02049	0,310	91,022	4,762
85	lunettes	0,01898	0,310	91,331	4,762
86	science	0,01898	0,310	91,641	4,762
87	stylo-bille	0,01898	0,310	91,950	4,762
88	TICE	0,01898	0,310	92,260	4,762
89	CD	0,01758	0,310	92,570	4,762
90	corbeille à papier	0,01758	0,310	92,879	4,762
91	encre indélébile	0,01758	0,310	93,189	4,762
92	géographie	0,01758	0,310	93,498	4,762

93	Internet	0,01758	0,310	93,808	4,762
94	stylo-feutre	0,01758	0,310	94,118	4,762
95	waters	0,01758	0,310	94,427	4,762
96	cour de récréation	0,01628	0,310	94,737	4,762
97	meuble	0,01628	0,310	95,046	4,762
98	chimie	0,01508	0,310	95,356	4,762
99	toit	0,01508	0,310	95,666	4,762
100	laboratoire	0,01397	0,310	95,975	4,762
101	plafond	0,01397	0,310	96,285	4,762
102	devoir	0,01293	0,310	96,594	4,762
103	dictionnaire	0,01293	0,310	96,904	4,762
104	clé	0,01198	0,310	97,214	4,762
105	lumière	0,01110	0,310	97,523	4,762
106	tableau noir	0,01110	0,310	97,833	4,762
107	ballon	0,01028	0,310	98,142	4,762
108	balle	0,00952	0,310	98,452	4,762
109	plumier	0,00952	0,310	98,762	4,762
110	année scolaire	0,00817	0,310	99,071	4,762
111	enseigner	0,00756	0,310	99,381	4,762
112	étudier	0,00700	0,310	99,690	4,762
113	surveillant	0,00557	0,310	100,000	4,762

3.2.5. La ville

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	rue	0,57446	3,856	3,856	71,429
2	voiture	0,54326	3,856	7,712	71,429
3	bâtiment	0,30404	2,057	9,769	38,095
4	route	0,29707	2,314	12,082	42,857
5	maison	0,29328	1,799	13,882	33,333
6	place	0,28357	2,057	15,938	38,095
7	mairie	0,26474	2,571	18,509	47,619
8	cinema	0,25159	2,571	21,080	47,619
9	avenue	0,24417	1,799	22,879	33,333
10	boulangerie	0,22319	2,571	25,450	47,619
11	magasin	0,21559	2,057	27,506	38,095
12	vélo	0,19323	2,057	29,563	38,095
13	autobús	0,19028	1,542	31,105	28,571
14	église	0,18193	2,571	33,676	47,619
15	pólice	0,17473	1,542	35,219	28,571
16	restaurant	0,17009	1,542	36,761	28,571
17	taxi	0,15581	1,028	37,789	19,048
18	métro	0,14976	1,799	39,589	33,333
19	train	0,14397	1,285	40,874	23,810
20	parc	0,13937	1,542	42,416	28,571
21	musée	0,13150	1,285	43,702	23,810
22	boucherie	0,12872	1,285	44,987	23,810
23	supermarché	0,12789	1,542	46,530	28,571
24	école	0,12605	1,542	48,072	28,571
25	jardín	0,12335	1,285	49,357	23,810
26	gare	0,11354	1,028	50,386	19,048
27	théâtre	0,11338	1,542	51,928	28,571
28	quartier	0,11321	1,028	52,956	19,048
29	fontaine	0,10618	1,028	53,985	19,048
30	hôtel de ville	0,10535	0,771	54,756	14,286
31	Carrefour	0,10427	1,028	55,784	19,048
32	appartement	0,10259	0,771	56,555	14,286
33	moto	0,10152	1,028	57,584	19,048
34	hôpital	0,09920	1,028	58,612	19,048
35	bar	0,09480	1,285	59,897	23,810
36	poste	0,09302	0,771	60,668	14,286
37	immeuble	0,09297	0,771	61,440	14,286
38	feu rouge	0,09212	1,028	62,468	19,048
39	ruelle	0,08892	0,771	63,239	14,286
40	feu	0,07441	0,514	63,753	9,524
41	pâtisserie	0,07272	1,285	65,039	23,810
42	édifice	0,07163	0,514	65,553	9,524
43	bistrot	0,07153	0,771	66,324	14,286
44	rond-point	0,07114	0,771	67,095	14,286
45	centre ville	0,07111	0,514	67,609	9,524

46	arbre	0,06882	1,028	68,638	19,048
47	boulevard	0,06451	0,514	69,152	9,524
48	camión	0,06445	0,514	69,666	9,524
49	discothèque	0,06310	0,771	70,437	14,286
50	aéroport	0,06245	0,771	71,208	14,286
51	marché	0,06201	0,771	71,979	14,286
52	centre	0,06165	0,514	72,494	9,524
53	rivière	0,05758	0,514	73,008	9,524
54	mer	0,05559	0,514	73,522	9,524
55	poissonnerie	0,05334	0,514	74,036	9,524
56	trottoir	0,05138	0,771	74,807	14,286
57	pollution	0,05039	0,514	75,321	9,524
58	panneau de signalisation	0,04889	0,514	75,835	9,524
59	cimetière	0,04878	0,514	76,350	9,524
60	bus	0,04835	0,514	76,864	9,524
61	banlieue	0,04818	0,514	77,378	9,524
62	arrêt de bus	0,04814	0,514	77,892	9,524
63	centre urbain	0,04762	0,257	78,149	4,762
64	lycée	0,04667	0,771	78,920	14,286
65	monument	0,04654	0,514	79,434	9,524
66	hotel	0,04331	0,514	79,949	9,524
67	crêperie	0,04124	0,257	80,206	4,762
68	Malaga	0,04124	0,257	80,463	4,762
69	pont	0,04124	0,257	80,720	4,762
70	boîte	0,04062	0,514	81,234	9,524
71	pub	0,04002	0,514	81,748	9,524
72	bureau de tabac	0,03838	0,257	82,005	4,762
73	centre-ville	0,03838	0,257	82,262	4,762
74	Lyon	0,03838	0,257	82,519	4,762
75	oiseau	0,03838	0,257	82,776	4,762
76	faubourg	0,03572	0,257	83,033	4,762
77	piéton	0,03572	0,257	83,290	4,762
78	collège	0,03546	0,514	83,805	9,524
79	marche	0,03324	0,257	84,062	4,762
80	plombier	0,03324	0,257	84,319	4,762
81	temps	0,03324	0,257	84,576	4,762
82	pharmacie	0,03274	0,514	85,090	9,524
83	usine	0,03128	0,514	85,604	9,524
84	fumée	0,03094	0,257	85,861	4,762
85	mosquée	0,03094	0,257	86,118	4,762
86	plage	0,03094	0,257	86,375	4,762
87	soleil	0,03094	0,257	86,632	4,762
88	stop	0,03094	0,257	86,889	4,762
89	vitrine	0,03094	0,257	87,147	4,762
90	cordonnier	0,03086	0,514	87,661	9,524
91	cathédrale	0,03002	0,514	88,175	9,524
92	station	0,02879	0,257	88,432	4,762
93	café	0,02872	0,514	88,946	9,524
94	banc	0,02680	0,257	89,203	4,762

95	fleuve	0,02680	0,257	89,460	4,762
96	garaje	0,02680	0,257	89,717	4,762
97	stress	0,02680	0,257	89,974	4,762
98	arène	0,02521	0,514	90,488	9,524
99	chien	0,02494	0,257	90,746	4,762
100	klaxon	0,02494	0,257	91,003	4,762
101	parking	0,02494	0,257	91,260	4,762
102	service social	0,02494	0,257	91,517	4,762
103	pavillon	0,02321	0,257	91,774	4,762
104	banque	0,02160	0,257	92,031	4,762
105	palace	0,02160	0,257	92,288	4,762
106	pompier	0,02160	0,257	92,545	4,762
107	port	0,02160	0,257	92,802	4,762
108	autoroute	0,02010	0,257	93,059	4,762
109	château	0,02010	0,257	93,316	4,762
110	palais	0,02010	0,257	93,573	4,762
111	bouchon	0,01871	0,257	93,830	4,762
112	cit��	0,01871	0,257	94,087	4,762
113	fleur	0,01871	0,257	94,344	4,762
114	tramway	0,01741	0,257	94,602	4,762
115	zone verte	0,01741	0,257	94,859	4,762
116	brasserie	0,01620	0,257	95,116	4,762
117	champ	0,01620	0,257	95,373	4,762
118	passage de pi��ton	0,01620	0,257	95,630	4,762
119	sortie	0,01620	0,257	95,887	4,762
120	universit��	0,01620	0,257	96,144	4,762
121	op��ra	0,01508	0,257	96,401	4,762
122	travailler	0,01403	0,257	96,658	4,762
123	culture	0,01306	0,257	96,915	4,762
124	gens	0,01306	0,257	97,172	4,762
125	d��partement	0,01215	0,257	97,429	4,762
126	Travail	0,01215	0,257	97,686	4,762
127	Bois (de Boulogne)	0,01131	0,257	97,943	4,762
128	transport	0,01131	0,257	98,201	4,762
129	adresse	0,01053	0,257	98,458	4,762
130	Courir	0,00980	0,257	98,715	4,762
131	lampadaire	0,00912	0,257	98,972	4,762
132	manger	0,00912	0,257	99,229	4,762
133	promenade	0,00848	0,257	99,486	4,762
134	F��te	0,00790	0,257	99,743	4,762
135	biblioth��que	0,00735	0,257	100,000	4,762

3.2.6. Moyens de transport

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	voiture	0,81069	8,434	8,434	100,000
2	vélo	0,73653	8,032	16,466	95,238
3	avion	0,70241	8,032	24,498	95,238
4	train	0,68864	7,229	31,727	85,714
5	bus	0,63406	6,426	38,153	76,190
6	métro	0,57888	6,827	44,980	80,952
7	bateau	0,55337	8,032	53,012	95,238
8	tram(way)	0,32880	4,418	57,430	52,381
9	cheval	0,30665	4,819	62,249	57,143
10	moto	0,29835	4,418	66,667	52,381
11	TGV	0,26937	3,213	69,880	38,095
12	taxi	0,23879	2,811	72,691	33,333
13	à pied	0,15119	2,008	74,699	23,810
14	RER	0,13012	1,606	76,305	19,048
15	TER	0,11373	1,205	77,510	14,286
16	camión	0,10494	1,205	78,715	14,286
17	motocyclette	0,09258	1,205	79,920	14,286
18	autobús	0,07768	0,803	80,723	9,524
19	parachute	0,07575	1,205	81,928	14,286
20	minibús	0,07261	1,205	83,133	14,286
21	scooter	0,06664	0,803	83,936	9,524
22	trottinette	0,06159	1,205	85,141	14,286
23	roller	0,05395	0,803	85,944	9,524
24	side-car	0,05055	0,803	86,747	9,524
25	hélicoptère	0,04777	0,803	87,550	9,524
26	barque	0,04249	1,205	88,755	14,286
27	motocycle	0,03246	0,402	89,157	4,762
28	car	0,03006	0,402	89,558	4,762
29	patin	0,03006	0,402	89,960	4,762
30	péniche	0,02793	0,803	90,763	9,524
31	funiculaire	0,02784	0,402	91,165	4,762
32	ómnibus	0,02784	0,402	91,566	4,762
33	patinette	0,02579	0,402	91,968	4,762
34	bagnole	0,02388	0,402	92,369	4,762
35	bicyclette	0,02388	0,402	92,771	4,762
36	chameau	0,02388	0,402	93,173	4,762
37	caisse	0,02212	0,402	93,574	4,762
38	charrette	0,02212	0,402	93,976	4,762
39	chemin de fer	0,02212	0,402	94,378	4,762
40	éléphant	0,02212	0,402	94,779	4,762
41	skate-board	0,02212	0,402	95,181	4,762
42	tracteur	0,02212	0,402	95,582	4,762
43	tanque	0,01898	0,402	95,984	4,762
44	AVE	0,01758	0,402	96,386	4,762
45	patin à roulettes	0,01758	0,402	96,787	4,762

46	sousmarin	0,01758	0,402	97,189	4,762
47	jet	0,01508	0,402	97,590	4,762
48	transatlantique	0,01508	0,402	97,992	4,762
49	fusée	0,01397	0,402	98,394	4,762
50	parapente	0,01293	0,402	98,795	4,762
51	ski	0,01198	0,402	99,197	4,762
52	traineau	0,01110	0,402	99,598	4,762
53	calèche	0,01028	0,402	100,000	4,762

3.2.7. Jeux et loisirs

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	football	0,59800	5,674	5,674	76,190
2	basket(-ball)	0,57890	5,674	11,348	76,190
3	cinema	0,55068	5,319	16,667	71,429
4	tennis	0,29060	3,901	20,567	52,381
5	jeu vidéo	0,25802	2,837	23,404	38,095
6	théâtre	0,25565	2,482	25,887	33,333
7	rugby	0,19964	1,773	27,660	23,810
8	volley-ball	0,17657	1,773	29,433	23,810
9	Monopoly	0,17422	1,773	31,206	23,810
10	natation	0,17264	2,482	33,688	33,333
11	sport	0,15193	1,773	35,461	23,810
12	Internet	0,14607	1,773	37,234	23,810
13	fête	0,14277	1,418	38,652	19,048
14	vélo	0,14125	2,128	40,780	28,571
15	musique	0,12769	1,064	41,844	14,286
16	poker	0,12635	1,418	43,262	19,048
17	danser	0,11618	1,773	45,035	23,810
18	discothèque	0,11414	1,418	46,454	19,048
19	jeu de l'oie	0,10303	1,064	47,518	14,286
20	musée	0,10076	1,418	48,936	19,048
21	base-ball	0,10058	1,418	50,355	19,048
22	Cluedo	0,09172	0,709	51,064	9,524
23	plage	0,08389	1,064	52,128	14,286
24	promener	0,07928	1,064	53,191	14,286
25	carte	0,07614	1,064	54,255	14,286
26	promenade	0,07589	0,709	54,965	9,524
27	randonnée	0,07412	1,064	56,028	14,286
28	restaurant	0,07288	0,709	56,738	9,524
29	Play Station	0,06989	0,709	57,447	9,524
30	Wii	0,06869	0,709	58,156	9,524
31	marelle	0,06568	0,709	58,865	9,524
32	pub	0,06568	0,709	59,574	9,524
33	Trivial	0,06510	0,709	60,284	9,524
34	ballon prisonnier	0,06498	1,064	61,348	14,286
35	café	0,06252	0,709	62,057	9,524
36	ski	0,06012	0,709	62,766	9,524
37	télévision	0,05996	0,709	63,475	9,524
38	faire la fête	0,05983	0,709	64,184	9,524
39	équitation	0,05634	0,709	64,894	9,524
40	lire	0,05553	0,709	65,603	9,524
41	padel	0,05459	1,064	66,667	14,286
42	tarot	0,05458	0,709	67,376	9,524
43	saut à la corde	0,05452	1,064	68,440	14,286
44	toupie	0,05395	0,709	69,149	9,524
45	sortir	0,05003	0,709	69,858	9,524

46	cap ou pas cap?	0,04762	0,355	70,213	4,762
47	Pro Evolution Soccer	0,04762	0,355	70,567	4,762
48	chat perché	0,04628	0,709	71,277	9,524
49	cache-cache	0,04601	0,709	71,986	9,524
50	squash	0,04514	0,709	72,695	9,524
51	Buzz!	0,04410	0,355	73,050	4,762
52	cirque	0,04410	0,355	73,404	4,762
53	lectura	0,04410	0,355	73,759	4,762
54	parc	0,04410	0,355	74,113	4,762
55	chasse	0,04085	0,355	74,468	4,762
56	dames	0,04085	0,355	74,823	4,762
57	loto	0,04085	0,355	75,177	4,762
58	ludo	0,04085	0,355	75,532	4,762
59	centre commercial	0,03783	0,355	75,887	4,762
60	loto sportif	0,03783	0,355	76,241	4,762
61	pêche	0,03783	0,355	76,596	4,762
62	pétanque	0,03783	0,355	76,950	4,762
63	roller	0,03783	0,355	77,305	4,762
64	opéra	0,03677	0,709	78,014	9,524
65	chanter	0,03504	0,355	78,369	4,762
66	échecs	0,03504	0,355	78,723	4,762
67	jeu de société	0,03504	0,355	79,078	4,762
68	quarté	0,03504	0,355	79,433	4,762
69	radio	0,03504	0,355	79,787	4,762
70	scrabble	0,03504	0,355	80,142	4,762
71	tourisme	0,03504	0,355	80,496	4,762
72	weekend	0,03504	0,355	80,851	4,762
73	bibliothèque	0,03246	0,355	81,206	4,762
74	poursuite	0,03246	0,355	81,560	4,762
75	samedi	0,03246	0,355	81,915	4,762
76	saut à l'élastique	0,03007	0,709	82,624	9,524
77	bilboquet	0,03006	0,355	82,979	4,762
78	gymnase	0,03006	0,355	83,333	4,762
79	naviguer sur Internet	0,03006	0,355	83,688	4,762
80	vendredi	0,03006	0,355	84,043	4,762
81	visiter la ville	0,03006	0,355	84,397	4,762
82	activité à l'air libre	0,02784	0,355	84,752	4,762
83	dimanche	0,02784	0,355	85,106	4,762
84	disco	0,02784	0,355	85,461	4,762
85	monument	0,02784	0,355	85,816	4,762
86	piscine	0,02784	0,355	86,170	4,762
87	Super Mario	0,02784	0,355	86,525	4,762
88	biscuit	0,02579	0,355	86,879	4,762
89	camping	0,02579	0,355	87,234	4,762
90	handball	0,02579	0,355	87,589	4,762
91	jeu de mots	0,02579	0,355	87,943	4,762
92	patinoire	0,02579	0,355	88,298	4,762
93	plongée	0,02579	0,355	88,652	4,762

94	alcohol	0,02388	0,355	89,007	4,762
95	bande dessinée	0,02388	0,355	89,362	4,762
96	Blokus	0,02388	0,355	89,716	4,762
97	course à pied	0,02388	0,355	90,071	4,762
98	faire du shopping	0,02388	0,355	90,426	4,762
99	alpinisme	0,02212	0,355	90,780	4,762
100	concert	0,02212	0,355	91,135	4,762
101	Facebook	0,02212	0,355	91,489	4,762
102	sauter	0,02212	0,355	91,844	4,762
103	crier	0,02049	0,355	92,199	4,762
104	aguagym	0,01898	0,355	92,553	4,762
105	livre	0,01898	0,355	92,908	4,762
106	s'amuser	0,01898	0,355	93,262	4,762
107	courir	0,01758	0,355	93,617	4,762
108	la tomate	0,01758	0,355	93,972	4,762
109	voyage	0,01758	0,355	94,326	4,762
110	yoga	0,01758	0,355	94,681	4,762
111	changer	0,01628	0,355	95,035	4,762
112	Pilates	0,01628	0,355	95,390	4,762
113	série télévisée	0,01628	0,355	95,745	4,762
114	sortir avec les amies	0,01628	0,355	96,099	4,762
115	voyager	0,01628	0,355	96,454	4,762
116	enfant	0,01508	0,355	96,809	4,762
117	spinning	0,01508	0,355	97,163	4,762
118	fitness	0,01397	0,355	97,518	4,762
119	petit	0,01397	0,355	97,872	4,762
120	école	0,01293	0,355	98,227	4,762
121	golf	0,01293	0,355	98,582	4,762
122	beach-volley	0,01198	0,355	98,936	4,762
123	récréation	0,01198	0,355	99,291	4,762
124	cricket	0,00952	0,355	99,645	4,762
125	balle	0,00756	0,355	100,000	4,762

3.2.8. Professions et métiers

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	professeur	0,91321	7,023	7,023	100,000
2	boulangier	0,47595	4,348	11,371	61,905
3	policier	0,36628	3,344	14,716	47,619
4	pompier	0,31421	3,010	17,726	42,857
5	cuisinier	0,30796	3,344	21,070	47,619
6	médecin	0,29994	3,344	24,415	47,619
7	boucher	0,20824	1,672	26,087	23,810
8	plombier	0,20543	2,007	28,094	28,571
9	docteur	0,18388	2,007	30,100	28,571
10	chef	0,17698	2,007	32,107	28,571
11	infirmier	0,16721	2,007	34,114	28,571
12	vendeur	0,16374	2,341	36,455	33,333
13	architecte	0,16295	1,672	38,127	23,810
14	traducteur	0,16170	2,341	40,468	33,333
15	pâtissier	0,15971	1,672	42,140	23,810
16	chauffeur	0,15826	1,672	43,813	23,810
17	dentiste	0,14670	1,672	45,485	23,810
18	maître	0,14665	1,338	46,823	19,048
19	serveur	0,14097	2,341	49,164	33,333
20	acteur	0,12432	1,672	50,836	23,810
21	jardinier	0,12093	1,003	51,839	14,286
22	avocat	0,10434	2,007	53,846	28,571
23	ingénieur	0,10072	1,338	55,184	19,048
24	barman	0,09965	1,338	56,522	19,048
25	pêcheur	0,09616	1,003	57,525	14,286
26	cordonnier	0,09590	1,003	58,528	14,286
27	charcutier	0,08495	0,669	59,197	9,524
28	informaticien	0,08495	0,669	59,866	9,524
29	garçon	0,08411	1,003	60,870	14,286
30	pharmacien	0,07797	1,003	61,873	14,286
31	gendarme	0,07656	0,669	62,542	9,524
32	militaire	0,06623	0,669	63,211	9,524
33	coiffeur	0,06510	0,669	63,880	9,524
34	charpentier	0,06459	0,669	64,548	9,524
35	directeur	0,06252	0,669	65,217	9,524
36	étudiant	0,06083	0,669	65,886	9,524
37	poissonnier	0,06030	0,669	66,555	9,524
38	concierge	0,05918	0,669	67,224	9,524
39	secrétaire	0,05832	0,669	67,893	9,524
40	interprète	0,05474	1,003	68,896	14,286
41	sommelier	0,05395	0,669	69,565	9,524
42	comptable	0,05173	0,669	70,234	9,524
43	facteur	0,04850	1,003	71,237	14,286
44	journaliste	0,04764	0,669	71,906	9,524
45	conducteur	0,04542	0,669	72,575	9,524

46	électricien	0,04444	0,669	73,244	9,524
47	commercial	0,04085	0,334	73,579	4,762
48	pédicure	0,04085	0,334	73,913	4,762
49	musicien	0,03970	0,669	74,582	9,524
50	fleuriste	0,03807	0,669	75,251	9,524
51	gardien de la paix	0,03795	0,669	75,920	9,524
52	politicien	0,03795	0,669	76,589	9,524
53	économiste	0,03783	0,334	76,923	4,762
54	employé de banque	0,03783	0,334	77,258	4,762
55	proctologue	0,03783	0,334	77,592	4,762
56	servante	0,03783	0,334	77,926	4,762
57	archéologue	0,03504	0,334	78,261	4,762
58	constructeur	0,03504	0,334	78,595	4,762
59	paysan	0,03504	0,334	78,930	4,762
60	psychologue	0,03504	0,334	79,264	4,762
61	styliste	0,03504	0,334	79,599	4,762
62	ouvrier	0,03445	0,669	80,268	9,524
63	juge	0,03386	0,669	80,936	9,524
64	agriculteur	0,03246	0,334	81,271	4,762
65	chef de cuisine	0,03006	0,334	81,605	4,762
66	fruitier	0,03006	0,334	81,940	4,762
67	routier	0,03006	0,334	82,274	4,762
68	tendeur	0,03006	0,334	82,609	4,762
69	biologiste	0,02801	0,669	83,278	9,524
70	broker	0,02784	0,334	83,612	4,762
71	chirurgien	0,02784	0,334	83,946	4,762
72	nettoyant	0,02784	0,334	84,281	4,762
73	sanitaire	0,02784	0,334	84,615	4,762
74	cassier	0,02579	0,334	84,950	4,762
75	chef de rang	0,02579	0,334	85,284	4,762
76	chevalier	0,02579	0,334	85,619	4,762
77	épicier	0,02579	0,334	85,953	4,762
78	hôte de vol	0,02579	0,334	86,288	4,762
79	officier	0,02579	0,334	86,622	4,762
80	papetier	0,02579	0,334	86,957	4,762
81	footballeur	0,02388	0,334	87,291	4,762
82	joueur	0,02388	0,334	87,625	4,762
83	maroquinier	0,02388	0,334	87,960	4,762
84	pilote	0,02388	0,334	88,294	4,762
85	valet	0,02388	0,334	88,629	4,762
86	vétérinaire	0,02388	0,334	88,963	4,762
87	voleur	0,02388	0,334	89,298	4,762
88	caviste	0,02212	0,334	89,632	4,762
89	chanteur	0,02212	0,334	89,967	4,762
90	fonctionnaire	0,02212	0,334	90,301	4,762
91	peintre	0,02212	0,334	90,635	4,762
92	postier	0,02212	0,334	90,970	4,762
93	retraité	0,02212	0,334	91,304	4,762
94	spécialiste	0,02212	0,334	91,639	4,762
95	banquier	0,02049	0,334	91,973	4,762

96	écrivain	0,02049	0,334	92,308	4,762
97	actrice	0,01898	0,334	92,642	4,762
98	libraire	0,01898	0,334	92,977	4,762
99	réceptionniste	0,01898	0,334	93,311	4,762
100	conseiller	0,01758	0,334	93,645	4,762
101	maçon	0,01758	0,334	93,980	4,762
102	récolteur	0,01758	0,334	94,314	4,762
103	videur	0,01758	0,334	94,649	4,762
104	navigateur	0,01628	0,334	94,983	4,762
105	barmaid	0,01508	0,334	95,318	4,762
106	marchand	0,01508	0,334	95,652	4,762
107	technicien	0,01508	0,334	95,987	4,762
108	employé	0,01397	0,334	96,321	4,762
109	surveillant	0,01397	0,334	96,656	4,762
110	chef d'entreprise	0,01293	0,334	96,990	4,762
111	femme de ménage	0,01293	0,334	97,324	4,762
112	opticien	0,01293	0,334	97,659	4,762
113	guichetier	0,01198	0,334	97,993	4,762
114	fermier	0,01110	0,334	98,328	4,762
115	pédagogue	0,01028	0,334	98,662	4,762
116	pilote d'avion	0,01028	0,334	98,997	4,762
117	hôtesse de l'air	0,00952	0,334	99,331	4,762
118	traumatologue	0,00952	0,334	99,666	4,762
119	gynécologue	0,00882	0,334	100,000	4,762

3.2.9. Ordinateurs et Internet

	Palabra	Disponibilidad	Frecuencia relativa	Frecuencia acumulada	% Aparición
1	souris	0,66379	5,079	5,079	76,190
2	clavier	0,60541	5,079	10,159	76,190
3	écran	0,51024	4,127	14,286	61,905
4	Internet	0,31311	2,540	16,825	38,095
5	Facebook	0,29120	3,492	20,317	52,381
6	page web	0,26384	2,222	22,540	33,333
7	Google	0,19372	2,540	25,079	38,095
8	logiciel	0,19338	1,905	26,984	28,571
9	clé USB	0,18674	1,905	28,889	28,571
10	ordinateur	0,17694	1,270	30,159	19,048
11	chat	0,16909	1,905	32,063	28,571
12	câble	0,15467	1,587	33,651	23,810
13	site (web)	0,14077	1,270	34,921	19,048
14	imprimante	0,14032	1,587	36,508	23,810
15	JClic	0,13667	1,905	38,413	28,571
16	CD	0,13585	1,270	39,683	19,048
17	chatter	0,12789	1,587	41,270	23,810
18	Hot Potatoes	0,12580	1,905	43,175	28,571
19	réseau social	0,11968	1,270	44,444	19,048
20	Wi-Fi	0,11470	1,270	45,714	19,048
21	hardware	0,11456	1,270	46,984	19,048
22	scanner	0,10656	1,587	48,571	23,810
23	Tuenti	0,10495	1,270	49,841	19,048
24	page	0,10337	0,952	50,794	14,286
25	Spotify	0,09721	1,270	52,063	19,048
26	courrier électronique	0,09493	1,270	53,333	19,048
27	(e-)mail	0,09415	1,270	54,603	19,048
28	touche	0,09325	0,952	55,556	14,286
29	virus	0,09265	1,587	57,143	23,810
30	software	0,08854	0,952	58,095	14,286
31	portable	0,08847	0,635	58,730	9,524
32	téléchargement	0,08545	0,635	59,365	9,524
33	jeu	0,08087	0,952	60,317	14,286
34	disquette	0,07768	0,635	60,952	9,524
35	programme	0,07589	0,635	61,587	9,524
36	web	0,07582	0,952	62,540	14,286
37	télécharger	0,07566	0,952	63,492	14,286
38	Twitter	0,07480	0,952	64,444	14,286
39	dossier	0,07092	0,952	65,397	14,286
40	Wikipedia	0,06785	0,952	66,349	14,286
41	(Microsoft) Word	0,06184	0,952	67,302	14,286
42	copier	0,06172	0,635	67,937	9,524
43	fórum	0,05794	0,952	68,889	14,286
44	coller	0,05716	0,635	69,524	9,524
45	matériel	0,05713	0,635	70,159	9,524

	informatique				
46	surfer	0,05634	0,635	70,794	9,524
47	DVD	0,05295	0,635	71,429	9,524
48	informatique	0,04762	0,317	71,746	4,762
49	(Microsoft)				
	PowerPoint	0,04666	0,952	72,698	14,286
50	Windows	0,04634	0,635	73,333	9,524
51	série	0,04601	0,635	73,968	9,524
52	antivirus	0,04438	0,952	74,921	14,286
53	enregistrer	0,04410	0,317	75,238	4,762
54	computer	0,04410	0,317	75,556	4,762
55	serveur	0,04410	0,317	75,873	4,762
56	webcam	0,04336	0,635	76,508	9,524
57	jeu vidéo	0,04085	0,317	76,825	4,762
58	effacer	0,04085	0,317	77,143	4,762
59	PSP	0,04085	0,317	77,460	4,762
60	écran LCD	0,03783	0,317	77,778	4,762
61	MP4	0,03783	0,317	78,095	4,762
62	netbook	0,03783	0,317	78,413	4,762
63	tour	0,03677	0,635	79,048	9,524
64	définition	0,03504	0,317	79,365	4,762
65	espace	0,03504	0,317	79,683	4,762
66	Hotmail	0,03504	0,317	80,000	4,762
67	vidéo	0,03279	0,635	80,635	9,524
68	USB	0,03246	0,317	80,952	4,762
69	réseau	0,03246	0,317	81,270	4,762
70	puce	0,03246	0,317	81,587	4,762
71	spam	0,03191	0,635	82,222	9,524
72	mémoire	0,03006	0,317	82,540	4,762
73	installer	0,03006	0,317	82,857	4,762
74	Youtube	0,03006	0,317	83,175	4,762
75	Adobe Photoshop	0,02784	0,317	83,492	4,762
76	ordinateur portable	0,02784	0,317	83,810	4,762
77	messagerie	0,02579	0,317	84,127	4,762
78	naviguer	0,02579	0,317	84,444	4,762
79	Microsoft Excel	0,02579	0,317	84,762	4,762
80	Wii	0,02579	0,317	85,079	4,762
81	moteur de recherche	0,02579	0,317	85,397	4,762
82	Busuu	0,02388	0,317	85,714	4,762
83	imprimeur	0,02388	0,317	86,032	4,762
84	MSN	0,02388	0,317	86,349	4,762
85	hi-fi	0,02212	0,317	86,667	4,762
86	opérateur	0,02212	0,317	86,984	4,762
87	Microsoft Works	0,02212	0,317	87,302	4,762
88	parler	0,02212	0,317	87,619	4,762
89	eMule	0,02049	0,317	87,937	4,762
90	navigateur	0,02049	0,317	88,254	4,762
91	Microsoft	0,02049	0,317	88,571	4,762
92	blog	0,02049	0,317	88,889	4,762

93	Ares	0,01898	0,317	89,206	4,762
94	courriel	0,01898	0,317	89,524	4,762
95	problème	0,01898	0,317	89,841	4,762
96	Yahoo	0,01898	0,317	90,159	4,762
97	Linux	0,01898	0,317	90,476	4,762
98	électromagnétique	0,01898	0,317	90,794	4,762
99	Les Sims	0,01758	0,317	91,111	4,762
100	adresse	0,01758	0,317	91,429	4,762
101	technicien	0,01758	0,317	91,746	4,762
102	système opératif	0,01758	0,317	92,063	4,762
103	connexion	0,01628	0,317	92,381	4,762
104	questionnaire à choix multiple	0,01628	0,317	92,698	4,762
105	poste de travail	0,01628	0,317	93,016	4,762
106	WordReference	0,01508	0,317	93,333	4,762
107	domaine	0,01508	0,317	93,651	4,762
108	graveur	0,01508	0,317	93,968	4,762
109	poubelle	0,01508	0,317	94,286	4,762
110	étude	0,01397	0,317	94,603	4,762
111	environnement	0,01397	0,317	94,921	4,762
112	disque dur	0,01397	0,317	95,238	4,762
113	musique	0,01397	0,317	95,556	4,762
114	journal en ligne	0,01397	0,317	95,873	4,762
115	recherche	0,01293	0,317	96,190	4,762
116	disque dur externe	0,01293	0,317	96,508	4,762
117	CityVille	0,01198	0,317	96,825	4,762
118	Microsoft Office	0,01198	0,317	97,143	4,762
119	film	0,01198	0,317	97,460	4,762
120	campus virtuel	0,01110	0,317	97,778	4,762
121	Skype	0,01110	0,317	98,095	4,762
122	interplex	0,01028	0,317	98,413	4,762
123	Web 2.0	0,00952	0,317	98,730	4,762
124	dictionnaire	0,00952	0,317	99,048	4,762
125	s'amuser	0,00882	0,317	99,365	4,762
126	extension	0,00649	0,317	99,683	4,762
127	image	0,00557	0,317	100,000	4,762